TEXT IS CROSS IN THE BOOK

B. G. TEUBNERS SAMMLUNG VON LEHRBÜCHERN
AUF DEM GEBIETE DER
MATHEMATISCHEN WISSENSCHAFTEN
MIT EINSCHLUSS IHRER ANWENDUNGEN
BAND XVII

WAHRSCHEINLICHKEITSRECHNUNG UND KOLLEKTIVMASSLEHRE

VON

DR. HEINRICH BRUNS
PROFESSOR DER ASTRONOMIE AN DER UNIVERSITÄT ZU LEIPZIG

LEIPZIG UND BERLIN
DRUCK UND VERLAG VON B. G. TEUBNER
1906

Vorwort.

Das vorliegende Buch ist aus den Vorlesungen entstanden, die ich seit fast 25 Jahren mit etwa zweijährigen Pausen über Wahrscheinlichkeitsrechnung gehalten habe, und es dürfte wohl kaum ein Nachteil sein, wenn sich diese Entstehungsweise in der Form der Darstellung deutlich verrät. Nachdem es mir gelungen war, eine brauchbare analytische Darstellung für willkürliche Verteilungsfunktionen aufzufinden, habe ich das Hauptgewicht auf die Entwicklung der Kollektivmaßlehre gelegt, die rund zwei Drittel des Werkes ausfüllt. Die sogenannten Anwendungen der Wahrscheinlichkeitsrechnung auf Versicherungswesen, Statistik und Fehlertheorie habe ich nur flüchtig gestreift, weil diese Gegenstände längst über den Rahmen einer bloßen Anwendung hinausgewachsen sind und eine selbständige Behandlung beanspruchen dürfen. Als Ausgleich bietet das Werk die erste lehrbuchmäßige Darstellung der allgemeinen Kollektivmaßlehre.

Das Manuskript war im unmittelbaren Anschluß an eine im Winter 1898/99 gehaltene Vorlesung niedergeschrieben worden und lag im Anfang des Jahres 1900 druckfertig vor, blieb jedoch einstweilen unveröffentlicht, weil ich mich damals mit dem Gedanken trug, eine größere Reihe von Anwendungen der Kollektivmaßlehre hinzuzufügen. Dieser Gedanke wurde jedoch später wieder aufgegeben, weil seine Ausführung einen Band für sich beansprucht hätte.

Bei der nochmaligen Durchsicht der vor sechs Jahren niedergeschriebenen Darstellung habe ich die Auffassung der Probleme und ihre Behandlung absichtlich ungeändert gelassen und mich darauf beschränkt, außer einigen Fußnoten den Abschnitt über das Gruppenschema und die Zahlenbeispiele für die Wirkung übertriebener Abrundung hinzuzufügen. Ich lege einigen Wert darauf, diese Tatsache hervorzuheben, weil in den letzten Jahren über einzelne Kapitel der Kollektivmaßlehre verschiedene Arbeiten erschienen sind, deren Inhalt sich mit der von mir gegebenen Darstellung sehr nahe berührt.

Leipzig, 1905 Juli 11.

H. Bruns.

XVII. Vorlesung.

Das gewöhnliche Urnensschema und seine Erweiterung.

XVIII. Vorlesung.

Schema für seltene Ereignisse und für Gruppen.

XIX. Vorlesung.

Der Bayessche Satz.

XX. Vorlesung.

Numerische Bearbeitung: direkte Mittelbildung.

XXI. Vorlesung.

Numerische Bearbeitung: Summenmethode (I. Form).

XXII. Vorlesung.

Numerische Bearbeitung: Summenmethode (II. Form).

XXIII. Vorlesung.

Numerische Bearbeitung: Beobachtungsgleichungen.

Berichtigungen.

S. 27, Z. 11 v. u.: statt letzten zu lesen beiden letzten.
S. 30, Z. 14 v. u.: statt $\mathfrak{W}$ zu lesen $\mathfrak{W}_v$.
S. 45, Formel (20): π zu streichen.
S. 122, Z. 19 v. o.: statt h zu lesen h_0.
S. 148, Z. 16 v. o.: statt S zu lesen $\mathfrak{S}$.
S. 166, Formel (5): statt s zu lesen s^2.
S. 168, Z. 12 v. u.: statt D zu lesen $\mathfrak{D}$.
S. 171, Z. 4 v. u.: m zu streichen.
S. 184, Z. 19 v. u.: Komma zu streichen.
S. 226, Z. 4 v. u.: statt $(1 - t^2)$ zu lesen $(1 - t)^2$.
S. 228, Z. 12 v. u.: die Exponenten s zu ersetzen durch k.
S. 303, Z. 3 v. u.: statt V zu lesen $\mathfrak{V}$.

Erste Vorlesung.

Gegenstand der Wahrscheinlichkeitsrechnung.

§ 1. Die Geschichte der Wahrscheinlichkeitsrechnung (kurz W.-R.) beginnt mit den Namen *Pascal* und *Fermat*. Ersterem wurden von dem *Chevalier de Méré* etliche auf die Glücksspiele jener Zeit bezügliche Aufgaben vorgelegt, von denen hier eine, die zur Klasse der sogenannten Teilungsprobleme gehört, besonders erwähnt werden soll. Der Gegenstand, um den es sich dabei handelte, war kurz folgender. Ein Glücksspiel zwischen den beiden Spielern A und B wird vor der Entscheidung abgebrochen; es soll der vorhandene Gesamteinsatz zwischen ihnen geteilt werden, und zwar nach Billigkeit, d. h. derart, daß der Spieler, der die größere Aussicht auf den Gewinn des ganzen Spieles hatte, auch einen entsprechend größeren Anteil erhält. Offenbar verlangt die Beantwortung der gestellten Frage, daß man die Gewinnaussichten von A und B in dem Augenblicke, wo das Spiel abgebrochen wurde, ziffermäßig und in einwandfreier Weise abzuschätzen vermag. *Pascal* löste die besondere, von *de Méré* gestellte Aufgabe, und ebenso auch *Fermat*, dem sie von *Pascal* vorgelegt wurde. Der Bericht hierüber findet sich teils in dem, allerdings nicht vollständig erhaltenen, Briefwechsel zwischen *Pascal* und *Fermat*[1]), teils in *Pascals „Traité du triangle arithmétique“*, der schon 1654 gedruckt vorlag, aber erst 1665, drei Jahre nach dem Tode *Pascals*, erschienen ist.

Es versteht sich von selbst, daß auch schon vor dem angeführten Ereignis die Menschen mit Wahrscheinlichkeiten gerechnet haben, in ähnlicher Weise, wie schon lange vor *Euklid* praktisch Geometrie getrieben worden ist. Gleichwohl ist man berechtigt, den Ursprung der W.-R. als einer mathematischen Disziplin von dem oben erwähnten Problem herzuleiten, denn zu seiner sachgemässen Behandlung mußten — und das ist die Hauptsache — zum ersten Male gewisse Begriffe und Grundsätze der W.-R. in mathematisch bestimmter Gestalt ausgesprochen werden. Der einmal gegebene Anstoß hat dann weiter gewirkt. Wenn auch die bald nachher erfolgte Auffindung der Infinitesimalrechnung und ihre Verbindung mit der Geometrie und Mechanik weiterhin vorzugsweise die Gedankenarbeit der Mathematiker in An-

1) Abgedruckt u. a. in Oeuvres de Pascal, édition Drion, Paris 1862, Vol. III.

spruch nahm, so finden wir doch bis zum Schlusse des achtzehnten Jahrhunderts in der Literatur der W.-R. fast alle die Namen vertreten, die während jener Periode in der Geschichte der Mathematik überhaupt hervorragen. Ebenso sind die Aufgaben, die auch heute noch den Hauptinhalt der W.-R. bilden, zum großen Teile bereits zu jener Zeit, wenigstens in den ersten Ansätzen, behandelt worden. Einen natürlichen Abschluß der genannten Periode bildet der *„Traité analytique des probabilités“* von *Laplace*, dessen erste Ausgabe 1812 erschien. Der „Traité“ enthält eine zusammenfassende Darstellung des Standes der W.-R., wie er sich zu jener Zeit, und zwar zum großen Teil durch die eigenen Arbeiten von *Laplace*, gestaltet hatte. Man darf sagen, daß seither, soweit es sich um die allgemeinen mathematischen Grundlagen der W.-R. handelt, nichts wesentlich neues hinzugekommen ist, während allerdings, wie zu erwarten war, die Anwendungen und die dafür geschaffenen besonderen Methoden eine stetige Weiterentwicklung erfahren haben. Hierher ist namentlich der zuerst von *Fechner* verfolgte Gedanke zu rechnen, daß die W.-R. zu einer Formenlehre der Kollektivgegenstände zu erweitern sei.

Die vorstehenden geschichtlichen Bemerkungen mögen hier genügen. Einen ausführlichen, biographisch geordneten Bericht findet man in dem Buche von *Todhunter „A history of the mathematical theory of probability from the time of Pascal to that of Laplace“ (Cambridge and London* 1865*)*.

§ 2. Unser nächster Schritt hätte jetzt darin zu bestehen, daß wir die allgemeinen Grundbegriffe der W.-R. entwickeln. Ehe wir jedoch dazu übergehen sind gewisse Vorfragen zu erledigen.

Die W.-R. wird häufig auch als die *mathematische Theorie der zufälligen Ereignisse* bezeichnet. Dieser Ausdruck deckt allerdings, wie wir später sehen werden, die Sache nicht ganz, denn einerseits ist nicht alles, was wir mit dem Namen Zufall belegen, dazu angetan, Gegenstand der W.-R. zu werden, andererseits ist die Geltung der Lehrsätze der W.-R. keineswegs auf die sogenannten zufälligen Ereignisse beschränkt. Wir wollen indessen für den Augenblick von diesen Besonderheiten absehen und uns daran halten, daß in den Darstellungen der W.-R. fortwährend von zufälligen Ereignissen die Rede ist. Dann wird man zunächst auf die Frage geführt, ob denn nicht etwa die Bezeichnung „Theorie des Zufalls“ ein Paradoxon, einen inneren Widerspruch enthalte, denn eine mathematische Theorie scheint doch nur da angebracht zu sein, wo es sich um bestimmte, nach Zahl und Maß erfaßbare Gesetzmäßigkeiten handelt, während der blinde Zufall ja gerade das Gegenteil von erkennbarer Gesetzmäßigkeit bedeutet. Hierauf läßt sich antworten, daß zufällige Ereignisse, eben so gut wie die erkennbar gesetzmäßigen, der Zählung oder Messung unterliegen können, und daß zwischen den Ergebnissen solcher Zählung

oder Messung Maßbeziehungen denkbar sind, die den Inhalt einer mathematischen Theorie zu bilden vermögen. Durch eine solche Antwort wird jedoch noch nicht die andere Frage erledigt, ob man denn bei einer wissenschaftlichen Betrachtungsweise überhaupt von einem blinden Zufall reden dürfe, und ob nicht etwa, wenn diese Frage zu verneinen ist, die mathematische Spekulation über die zufälligen Ereignisse lediglich auf ein müßiges Spiel des Verstandes hinauslaufe, dem keinerlei Bedeutung für die Vorgänge der Wirklichkeit zukommt.

Es kann nun nicht zweifelhaft sein, daß der Gang, den die Erziehung unseres Denkens im Laufe der Jahrhunderte genommen hat, mit der Zulassung eines blind waltenden Zufalls schlechterdings unverträglich ist. Der Geograph *Oskar Peschel* sagt einmal in seiner „Völkerkunde“ (zweite Auflage, Seite 399), da wo er auf die chinesische Kultur zu sprechen kommt, sehr bezeichnend folgendes: „Seit unserem geistigen Erwachen, seit wir als Mehrer der Kulturschätze aufgetreten sind, haben wir unverdrossen, mit den Schweißperlen auf der Stirn, nur nach einem Dinge gesucht, von dessen Dasein die Chinesen keine Ahnung haben, und für das sie auch schwerlich eine Schüssel Reis geben würden. Dieses eine unsichtbare Ding nennen wir Kausalität. An den Chinesen haben wir eine ungezählte Menge von Erfindungen bewundert und von ihnen uns angeeignet, aber wir verdanken ihnen nicht eine einzige Theorie, nicht einen einzigen tieferen Blick in den Zusammenhang und die nächsten Ursachen der Erscheinungen.“ Anders ausgedrückt können wir sagen, daß unser Denken sich erst dann befriedigt fühlt, wenn wir bei der Betrachtung eines Gegenstandes Ordnung und Zusammenhang hergestellt und seinen zureichenden Grund aufgedeckt haben; die Heranziehung eines blinden Zufalls gilt als Eingeständnis, daß wir diese Anforderungen nicht zu erfüllen vermögen. Das Verlangen nach einer zureichenden Begründung wird von uns auch nicht bloß bei der Betrachtung natürlicher Vorgänge erhoben, denn wir fordern z. B. von einem Drama, wenn es als Kunstwerk gelten soll, neben einer vollendeten Sprache vor allem eine überzeugende psychologische Motivierung der dargestellten Handlungen.

§ 3. Wenn nun, entsprechend den vorstehenden Bemerkungen, die Welt als ein Ganzes angesehen wird, in dem alles mit jedem und jedes mit allem, sei es mittelbar, sei es unmittelbar, zusammenhängt, so entsteht die Frage, wie mit dieser Auffassung der Umstand zu reimen sei, daß eine ausgebildete mathematische Theorie des Zufalls nicht bloß existiert, sondern auch mit Erfolg auf die Vorgänge der Wirklichkeit angewendet wird. Die Lösung dieses Widerspruchs ergibt sich aus folgender Bemerkung. Selbst wenn man die Zulassung eines blinden Zufalls grundsätzlich ablehnt, so bleibt doch die Tatsache bestehen, daß wir auf Schritt und Tritt Dingen und Vorgängen begegnen, die für uns in ausgesprochener Weise den *Charakter*

der Zufälligkeit besitzen. So sind wir beim Würfelspiel außer Stande, über den Erfolg des einzelnen Wurfes etwas bestimmtes vorherzusagen, ebenso vermögen wir nicht, zwischen den Ergebnissen verschiedener Würfe einen Zusammenhang zu erkennen, vielmehr erscheint uns der ganze Verlauf des Spieles als etwas, das durchaus dem blinden Zufall unterworfen ist. Das kommt denn auch in dem gewöhnlichen Sprachgebrauch zum Ausdruck, der solche Spiele mit dem Namen reiner Zufallsspiele belegt.

Die erwähnten, mit dem Charakter der Zufälligkeit behafteten oder *scheinbar* zufälligen Ereignisse sind es nun, die den Anlaß zur Entstehung und Ausbildung der W.-R. gegeben haben. Der Gedankengang, der dabei eingeschlagen wird, ist in Kürze folgender. Man konstruiert zunächst die *Vorstellung* von blind zufälligen Ereignissen, indem man bei den scheinbar zufälligen Ereignissen der Wirklichkeit vollständig von den kausalen Beziehungen abstrahiert, die tatsächlich, wenn auch für uns nicht erfaßbar, vorhanden sind. Diese Vorstellung ist, eben ihrer Entstehungsweise halber, eine reine Abstraktion, d. h. ein Gedankending, das so, wie es vorgestellt wird, nirgends Verwirklichung findet. Dadurch werden wir aber keineswegs gehindert, die nur gedachten blind zufälligen Ereignisse zum Gegenstande der mathematischen Untersuchung zu machen. Das Ergebnis einer solchen Untersuchung ist, wie der Erfolg gezeigt hat, ein selbständiges Kapitel der Mathematik, oder m. a. W. ein vollständiges Lehrgebäude mit eigenen Grundbegriffen und Methoden. Da es sich nun bei der Aufrichtung dieses Gebäudes immer nur um logische, an rein abstrakten Vorstellungen ausgeführte Operationen handelt, so besitzen die Lehrsätze der Zufallstheorie denselben Grad von Gewißheit, der irgend welchen, aus richtigen Vordersätzen auf richtige Weise gezogenen Schlüssen zuzusprechen ist. Das Gegenteil davon könnte nur dann eintreten, wenn die Analyse der Grundbegriffe der W.-R. in diesen innere Widersprüche aufdeckte; ein solcher Fall ist indessen, wie wir später erkennen werden, nicht zu befürchten.

Die Gewißheit der in der Zufallstheorie hergeleiteten Lehrsätze besteht zunächst nur für das Gebiet der bloß vorgestellten zufälligen Ereignisse. Nun deckt sich aber diese Vorstellung nach dem Obigen keineswegs mit der Beschaffenheit der scheinbar zufälligen Ereignisse der Wirklichkeit, und daraus fliessen sofort zwei Folgerungen.

Erstens nämlich muß man darauf verzichten, durch rein logische Deduktionen den Satz erhärten zu wollen, daß die Lehrsätze der W.-R. eine notwendige Geltung für die Vorgänge der Wirklichkeit besitzen; vielmehr läßt sich der Nachweis, daß jene Sätze zur Untersuchung beobachteter Ereignisse nun auch wirklich brauchbar seien, einzig und allein aus der Erfahrung führen.

Zweitens ersieht man, daß den Lehrsätzen der W.-R., soweit sie

sich in einem gegebenen Falle als brauchbar herausstellen, immer nur die Bedeutung einer Annäherung zuzusprechen ist.

Es ist wichtig, diese Punkte ausdrücklich hervorzuheben, denn ihre Nichtbeachtung hat manchen Mißbrauch der W.-R. verschuldet.

Nach den vorstehenden Bemerkungen ist also der oben hervorgehobene Widerspruch zwischen dem Kausalitätsprinzip und dem Bestehen einer tatsächlich brauchbaren Zufallstheorie nur scheinbar, denn die blind zufälligen Ereignisse der W.-R. sind nichts anderes, als eine für den Zweck der Untersuchung aufgestellte Voraussetzung, oder, wie man kurz zu sagen pflegt, eine Arbeitshypothese. *Der wissenschaftliche Wert dieser Hypothese ist darin zu suchen, daß ihre Konsequenzen, auf Gegenstände der Erfahrung angewendet, als Reagens auf das Vorhandensein oder Fehlen eines erkennbaren kausalen Zusammenhanges dienen.*

§ 4. Das hier charakterisierte Verfahren zur Aufrichtung eines Systems der W.-R. ist seinem Wesen nach nicht von dem Wege verschieden, den man überall da einschlägt, wo natürliche Vorgänge einer „exakten“ Behandlung unterworfen werden sollen. Wenn der Kosmos eine Einheit bildet, innerhalb deren jedes mit jedem zusammenhängt, so ist jeder Vorgang, auch der anscheinend einfachste, in Wahrheit unermeßlich verwickelt. Man muß also, da ja die Fähigkeiten unseres Verstandes begrenzt sind, im voraus darauf verzichten, einen solchen Vorgang vollständig zu erfassen. Diese vollständige Erfassung ist aber auch nicht erforderlich. Denn die Hilfsmittel unserer Wahrnehmung sind ebenfalls begrenzt, und es sind bei jedem Vorgange unermeßlich viele Umstände vorhanden, die unserer Wahrnehmung eben wegen ihrer Begrenztheit weder mittelbar noch unmittelbar zugänglich werden und die darum in einem gegebenen Falle als nicht vorhanden betrachtet werden dürfen, auch wenn uns ihre Existenz aus anderen Gründen gewiß ist. Wenn sich z. B. eine Bakterie in ihrer Nährlösung bewegt, so verursacht sie dadurch eine Änderung in der Anordnung der unserem Erdkörper zugerechneten Massen, und diese Änderung macht sich fühlbar nicht nur in der Bewegung des Erdkörpers und des Sonnensystems, sondern auch darüber hinaus bis zu dem fernsten Fixstern hin. Es würde gar keine Schwierigkeit bieten, den Betrag einer solchen Wirkung in den gebräuchlichen astronomischen Einheiten ziffernmäßig auszudrücken; auch ist es, wenn es sich nur um die Frage nach dem Vorhandensein einer Ursache handelt, ganz gleichgültig, ob der für die Größe der Wirkung anzusetzende Dezimalbruch hinter dem Komma mit nur fünf oder mit fünftausend oder mit fünfmillionen Nullen beginnt. Diese Nullenanzahl gewinnt erst dann Bedeutung, wenn es darauf ankommt, ob die betrachtete Größe bei der gerade vorgelegten Untersuchung als merklich oder unmerklich anzusehen ist.

Unter den hervorgehobenen Umständen sind wir also einerseits gezwungen, andererseits berechtigt, bei der Betrachtung der natürlichen Dinge von Abstraktionen, d. h. von vereinfachenden Voraussetzungen auszugehen. Allerdings wird diese Erleichterung, die den Versuch einer wissenschaftlichen Behandlung bestimmter Vorgänge überhaupt erst ausführbar macht, dadurch erkauft, daß die so zu erlangenden Resultate, der Wirklichkeit gegenüber, stets nur Annäherungen, d. h. notwendig unvollständig sind. Dabei sind die Schlüsse die von den zugrunde gelegten Abstraktionen theoretisch zu den Annäherungen führen, logisch völlig sicher, oder können es wenigstens sein, während die Prüfung, wie weit jene Schlüsse für die Einsicht in das Wesen beobachteter Vorgänge ausreichen, stets der Erfahrung vorbehalten bleibt. Es ist hier nicht nötig, diese Verhältnisse überall im einzelnen nachzuweisen; einige Beispiele werden genügen. So sind die vollkommen starren und die vollkommen elastischen Körper, mit denen man sich in der Mechanik beschäftigt, lediglich Abstraktionen, die nirgends in der Natur strenge Verwirklichung finden. Wenn auch diese Vorstellungen für das Entwerfen und das Verstehen zahlloser Mechanismen ausreichen, so ist doch der Ingenieur, der eine eiserne Brücke plant, genötigt, auf die Unvollkommenheit jener Abstraktionen ernstlich Rücksicht zu nehmen. Andere Beispiele sind die Lichtstrahlen der *Newton*schen Optik und die „reinen“ Substanzen des Chemikers. Im Anschlusse hieran kann man geradezu den paradox klingenden Satz aussprechen, daß in den messenden Naturwissenschaften die „Exaktheit“ in der Darstellung beobachteter Vorgänge um so größer ausfällt, je unexakter, d. h. je einfacher, die zugrunde gelegten Abstraktionen sein dürfen; es ist durchaus kein bloßer Zufall, daß unter den Naturwissenschaften die Astronomie die älteste, und die physikalische Meteorologie eine der jüngsten ist.

In den vorstehenden Bemerkungen haben wir die Übereinstimmung betont, die zwischen dem Vorgehen in der W.-R. und in anderen Theorien besteht. Daneben ist nun aber noch eine ausgeprägte Verschiedenheit vorhanden. Während nämlich überall sonst die Entwicklung einer Theorie darauf ausgeht, uns eine Einsicht in noch unerkannte Zusammenhänge zu verschaffen, so beginnt man in der W.-R. damit, gewisse unerkannte, aber sicher vorhandene Beziehungen einfach als nicht vorhanden zu setzen. Dieser Umstand verleiht offenbar der W.-R. eine ganz bestimmte Sonderstellung gegenüber allen übrigen Teilen der angewandten Mathematik.

§ 5. Die vorstehenden Bemerkungen führen naturgemäß zu der weiteren Frage, von welchen Umständen denn im gegebenen Falle der Charakter der Zufälligkeit eines Ereignisses abhängt. Daß es sich dabei, allgemein gesprochen, um einen Mangel unserer Erkenntnis handelt, geht zur Genüge aus der bisherigen Erörterung hervor, jedoch

lassen sich im einzelnen gewisse Unterschiede nachweisen. Vergleicht man z. B. die Blattformen unserer Bäume miteinander, so besitzen diese Formen für uns den Charakter der Zufälligkeit, denn es fehlt uns jede Einsicht in die Ursachen, die dem Blatte der Eiche, der Buche, des Ahorn usw. ihre wohlbekannte typische Gestalt gegeben haben. Hier ist also der „Zufall" eine Folge unserer vollständigen Unkenntnis des ursächlichen Zusammenhanges. Im Gegensatze dazu gibt es nun aber auch Vorgänge, bei denen wir die wirksamen Umstände, wenigstens in qualitativer Hinsicht, sehr vollständig übersehen und dennoch wiederum von Zufall sprechen, und zwar deshalb, weil der einzelne beobachtete Vorgang so verwickelt ist, daß wir außer Stande sind, ihn ziffernmäßig zu verfolgen. Man denke sich z. B. eine Lotterie in der Weise eingerichtet, daß in das Glücksrad statt der zusammengefalteten Loszettel numerierte Kugeln in ganz bestimmter Anordnung hineingelegt werden, ferner sei mit dem Rade ein Mechanismus verbunden, der jedesmal nach einer gewissen, genügend großen Zahl von Umdrehungen des Rades je eine Kugel austreten läßt, endlich werde das Ganze nicht durch Menschenhand, sondern durch ein aufgezogenes Triebwerk in Tätigkeit gesetzt, dann hat man es bei der Nummernfolge der austretenden Kugeln sicherlich mit einem Vorgange von streng mechanischer Gesetzmäßigkeit zu tun. Trotzdem wird niemand Bedenken tragen, jene Nummernfolge für ebenso zufällig anzusehen, wie etwa den Verlauf eines Würfelspiels. Hier liegt der Ursprung der Zufälligkeit offenbar nicht in unserer Unwissenheit über die wirksamen Ursachen, sondern wesentlich in der außerordentlichen Verwickeltheit des ganzen Vorganges.

Es verdient bemerkt zu werden, daß die zuletzt angeführte Entstehungsweise der Zufälligkeit eines Ereignisses schon von *Kepler* mit voller Bestimmtheit hervorgehoben worden ist. In der Schrift „*De stella nova in pede Serpentarii*" (Cap. XXVII) kommt *Kepler* auf den Zufall und das Würfelspiel zu sprechen und sagt: „*Quare hoc jactu Venus*[1]) *cecidit, illo canis? Nimirum lusor hac vice tesselam alio latere arripuit, aliter manu condidit, aliter intus agitavit, alio impetu animi manusve projecit, aliter interflavit aura, alio loco alvei impegit. Nihil hic est, quod sua causa caruerit, si quis ista subtilia posset consectari.*" Der letzte, hier besonders hervorgehobene Satz drückt offenbar genau das aus, was vorhin als Verwickeltheit eines im übrigen bekannten kausalen Zusammenhanges bezeichnet wurde.

Die vorstehend besprochenen Fälle, nämlich „Zufall wegen völliger Unwissenheit" und „Zufall wegen übergroßer Verwickelung", können als die beiden Extreme angesehen werden, zwischen denen sich, soweit

---

1) Venus und canis bezeichnen den besten und den schlechtesten Wurf bei dem alten Astragalspiel.

physische Dinge und Vorgänge in Frage kommen, alle anderen Fälle mit stetigen Übergängen einordnen. Daneben gibt es aber auch noch Fälle, bei denen die Verwickeltheit des ursächlichen Zusammenhanges den Anschein der Zufälligkeit erzeugt, obgleich wir diesen Zusammenhang vollständig und ziffernmäßig verfolgen können. Zur Erläuterung möge folgendes Beispiel dienen, auf das wir gelegentlich zurückzugreifen haben werden. Das große Tafelwerk von *Vega*, der *Thesaurus logarithmorum completus* enthält u. a. eine zehnstellige Tafel der gemeinen Logarithmen aller fünfziffrigen Zahlen, und zwar befinden sich auf jeder Seite fünf Spalten mit je sechzig Logarithmen. Zählt man nun in jeder Spalte ab, wie oft eine „Endnull", d. h. eine Null in der zehnten Decimale, auftritt, und stellt man ferner, etwa für die tausend ersten Spalten, eine Tabelle zusammen, die die Anzahl der Spalten mit 0, 1, 2, ... 60 Endnullen angibt, so findet man, daß die hierbei zum Vorschein kommende Verteilung der Endnullen in ihrer äusseren Gestalt einen ähnlichen Verlauf zeigt, wie er uns später bei Ereignissen von ausgesprochen zufälligem Charakter begegnen wird. Ein solches Verhalten ist offenbar deswegen bemerkenswert, weil im vorliegenden Falle der ursächliche Zusammenhang, nämlich die arithmetische Beziehung zwischen Numerus und Logarithmus, vollkommen offen daliegt und auch bis in jede Einzelheit hinein ziffernmäßig verfolgt werden kann, so daß alles Zufällige von vornherein ausgeschlossen ist.

Aus den angeführten Bemerkungen ist natürlich nicht der Schluß zu ziehen, daß, wenn ein verwickelter Zusammenhang vorliegt, nun auch das Ergebnis jedesmal den Charakter der Zufälligkeit tragen müsse. Wird z. B. ein Pulver, dessen Körnchen verschiedene Größe besitzen, auf bekannte Weise in einer Reihe hintereinander geschalteter Schlämmtrichter der Wirkung eines konstanten Wasserstromes ausgesetzt, so ist die Bewegung der einzelnen Körnchen sicherlich recht verwickelt, dagegen ist das Schlußergebnis des ganzen Vorganges, nämlich die Zerlegung des Pulvers nach der Korngröße, eine einfache und im voraus zu übersehende Sache. Das Gleiche gilt, wenn zwei verschiedene Pulver in dasselbe Gefäß geschüttet und durch geeignete Bewegungen des Gefäßes innig gemischt werden.

§ 6. Wenn der Zufall wesentlich daher rührt, daß wir von den Entstehungsgründen eines Dinges nichts oder fast nichts wissen, so vermag selbstverständlich keine Kunst des Mathematikers aus einem solchen Nichts eine Aussage von Belang herauszuholen. Daraus ist zu entnehmen, daß für die W.-R. wesentlich nur derjenige Zufall in Betracht kommt, der aus übergroßer Verwickeltheit entspringt. Man kann nun fragen, wie hierbei der Charakter der Zufälligkeit zustande kommt, da ja dieser nach den vorhergehenden Bemerkungen nicht notwendig aus der Verwickeltheit eines Vorganges folgt. Eine scharf-

sinnige Untersuchung darüber hat *J. von Kries* in einer bisher viel zu wenig beachteten kritischen Studie über die Prinzipien, auf denen die Anwendungen der W.-R. beruhen, gegeben.[1]) Es wird hier genügen, seinen Gedankengang an einem Beispiele deutlich zu machen.

Wir betrachten einen Würfel in dem Augenblicke, wo sein freier Fall beginnt. Die vollständige Bestimmung der Lage und der Bewegung des Würfels in dem gedachten Augenblicke erfordert, wie die Mechanik lehrt, die Kenntnis der numerischen Werte von zwölf Bestimmungsstücken oder Parametern, die wir uns auf irgend eine passende Art eingeführt denken und mit $x, x', x'', \ldots$ bezeichnen wollen. Die Gesamtheit der bei dem Spiel in Betracht zu ziehenden Wertsysteme der Parameter bildet eine gewisse zwölffach ausgedehnte Mannigfaltigkeit, und jede Stelle innerhalb dieser Mannigfaltigkeit entspricht einem bestimmten und möglichen Beginne des freien Falls. Die geworfene Augenzahl y hängt von den x ab, ist also, da y nur sechs verschiedene Werte annehmen kann, eine gewisse unstetige Funktion der x. Denkt man sich weiter die Umstände, die sonst noch auf den Verlauf des Wurfes einwirken, wie z. B. die physische Beschaffenheit von Würfel und Tisch, die Elastizität, die Reibungen usw., der Einfachheit halber frei von zeitlichen Änderungen, so ist die Gestalt der Funktion y völlig bestimmt. Ferner läßt sich, wenn wir auch nicht den analytischen Ausdruck für y aufzustellen vermögen, doch eine, hier wesentliche, Eigenschaft dieses Ausdruckes aussagen. Bedeutet nämlich b den Wert von y, der zu einem beliebig vorgeschriebenen Wertsystem $a, a', a'', \ldots$ der Parameter gehört, so läßt sich ein von b verschiedenes y stets durch eine äußerst winzige Änderung des Wertsystems der a herbeiführen. Die Funktion y ist also, wie wir kurz sagen wollen, äußerst rasch veränderlich. Damit wird dem Spieler die Möglichkeit genommen, das Ergebnis eines Wurfes vorherzusehen, ja er ist nicht einmal imstande, das Ergebnis eines soeben ausgeführten Wurfes absichtlich dadurch zu wiederholen, daß er sich bemüht, den Würfel genau in derselben Weise wie vorher zu werfen. Die eigentliche Quelle des Zufalls ist also hier in der raschen Veränderlichkeit von y zu suchen. Die Bedeutung dieses Umstandes tritt noch klarer hervor, wenn man sich die obigen Betrachtungen für ein Geschicklichkeitsspiel, wie z. B. Billard, wiederholt denkt.

Nach diesen Erörterungen wenden wir uns jetzt zur Entwicklung der Begriffe und Lehrsätze, auf denen sich die W.-R. aufbaut.

---

1) Die Prinzipien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung von *Joh. von Kries.* Freiburg i. B. 1886.

Zweite Vorlesung.

Grundbegriffe der Wahrscheinlichkeitsrechnung.

§ 7. Der gewöhnliche Sprachgebrauch versteht unter Wahrscheinlichkeit eines Ereignisses den Grund, den wir für die Annahme haben, daß das Ereignis eintreten werde oder eingetreten sei. Die Stärke dieses Grundes läßt Unterschiede zu, denn man gebraucht ohne weiteres die Ausdrucksweise, daß ein gewisses Ereignis wahrscheinlicher sei, als ein anderes. Wenn z. B. ein gewöhnlicher Würfel auf fünf Seiten eine Null, auf der sechsten aber eine Eins trägt, so besinnt sich niemand, dem Werfen der Null die größere Wahrscheinlichkeit zuzusprechen. Hierdurch wird man zu der Frage geführt, ob das „größer" und „kleiner" einer Wahrscheinlichkeit lediglich als ein Unterschied des Grades aufzufassen sei, oder aber eine Maßbestimmung, d. h. eine ziffernmäßige Abschätzung, zulasse. Wäre z. B. ersteres der Fall, so müßte man darauf verzichten, die Wahrscheinlichkeit von Ereignissen zum Gegenstande der Rechnung zu machen, denn dazu gehören notwendig Zahlenwerte. Die Sache läge dann ähnlich, wie bei den Begriffen schön, gut, weise usw., bei denen man ebenfalls ein „mehr" und ein „weniger" anerkennt, aber keine Maßbestimmungen unternimmt. In der W.-R. ist nun mit Erfolg der Versuch unternommen worden, ein Wahrscheinlichkeitsmaß aufzustellen, wobei jedoch schon hier bemerkt werden möge, daß die Zulässigkeit oder Brauchbarkeit dieses Maßes an gewisse Voraussetzungen gebunden ist, die nicht immer, wenn wir von Wahrscheinlichkeit schlechtweg sprechen, erfüllt zu sein brauchen. Um das Wesen der gedachten Maßbestimmung darzulegen, schicke ich folgende Bemerkung voraus.

Es sei eine aus c Gliedern bestehende Reihe von Dingen gegeben, die vorläufig als gleichberechtigt unterschiedslos nebeneinander gestellt werden. Diese Reihe werde nach irgend welchen Gesichtspunkten in die beiden einander ausschließenden Teilreihen E und F mit den Gliederzahlen a und b zerlegt; dann gibt nach einem gebräuchlichen Ausdrucke der Quotient $a:c$ oder $b:c$ die *„relative Häufigkeit"* (kurz rH., Mehrheit rHH.) der Dinge an, die in der betrachteten Reihe der Gruppe E oder F angehören. Denkt man sich weiter eine vorgelegte Menge gleichberechtigter Dinge auf verschiedene Arten in Teilmengen zerlegt und dazu die rHH. berechnet, so werden zwischen den letzteren gewisse, von der vorgenommenen Zerlegung abhängende Beziehungen stattfinden. Aus solchen Beziehungen läßt sich dann, falls ihre Zahl und Mannigfaltigkeit hinreichend groß ist, ein ganzes Lehrgebäude, nämlich eine *Häufigkeitsrechnung* errichten, und zwar in ähnlicher Weise, wie sich z. B. aus dem Begriffe der Teilbarkeit ganzer Zahlen eine „Zahlentheorie" entwickelt hat.

§ 8. Dies vorausgeschickt denken wir uns jetzt eine Urne, die a weiße und b schwarze Kugeln enthält, während die gesamte Füllung $a+b$ mit c bezeichnet werden soll. Die Kugeln seien durch fortlaufende Nummern unterschieden, im übrigen aber, d. h. abgesehen von der Farbe, einander gleich. Nachdem der Inhalt der Urne gehörig gemischt worden ist, werde eine Kugel blindlings gezogen. Hierbei soll der herkömmliche Ausdruck „blindlings“ zweierlei besagen: erstens nämlich, daß der Ziehende außer Stande sei, bewußt oder absichtlich eine bestimmte Farbe oder Kugel zu bevorzugen, und zweitens, daß für jede der c Kugeln die gleiche Möglichkeit des Gezogenwerdens bestehe, so daß der Zug selber den Charakter eines zufälligen Ereignisses annimmt.

Bezeichnet man weiter einen weißen Zug kurz als Ereignis E, so liefern von den c gleichmöglichen Zügen a das Ereignis E oder sind — nach der herkömmlichen Ausdrucksweise — für E „günstig“. Der Quotient $a:c$ liefert dann die rH. der für E günstigen Fälle. *Man ist nun übereingekommen,* diesen Quotienten in dem vorliegenden Falle als Maß der Wahrscheinlichkeit eines weißen Zuges zu wählen und ihn die *mathematische Wahrscheinlichkeit des Ereignisses E* zu nennen.

Es bedarf keiner weitläufigen Ausführungen, wie die vorstehend an einem Beispiele erklärte Definition auf andere Vorgänge zufälliger Art zu übertragen ist. Wenn bei einem Versuche, dessen Ergebnis dem Zufall unterworfen ist, das Ereignis E oder sein Gegenteil Nicht-E eintreten kann, so hat man zunächst die Reihe der denkbaren Fälle aufzustellen, und zwar in einer solchen Gestalt, daß die einzelnen Glieder voneinander unabhängig sind, ferner sich gegenseitig ausschließen und endlich die gleiche Möglichkeit des Eintretens besitzen. Dann hat man durch Abzählung die Menge der für E günstigen Fälle zu bestimmen, daraus ihre rH. zu berechnen und erhält hiermit sofort die gesuchte mathematische Wahrscheinlichkeit von E.

Da man es in der W.-R. immer nur mit mathematischen Wahrscheinlichkeiten zu tun hat, so werden wir weiterhin das Beiwort „mathematisch“ einfach fortlassen und schlechtweg von Wahrscheinlichkeit (kurz $\mathfrak{W}$., Mehrheit $\mathfrak{W}\mathfrak{W}$.) sprechen. Ferner werden wir öfters für die $\mathfrak{W}$. eines Ereignisses E das unmittelbar verständliche Zeichen $\mathfrak{W}(E)$ benutzen.

§ 9. Aus der aufgestellten Definition fließen sofort gewisse Folgerungen, die wir zunächst erörtern wollen. Da die Größe $\mathfrak{W}(E)$ ihrer Bedeutung nach nichts anderes, als die rH. der für E günstigen Fälle ist, und da ferner auf diesem Begriffe sich die ganze Zufallstheorie aufbaut, *so ist die W.-R. ihrem mathematischen Inhalte nach lediglich eine Häufigkeitsrechnung.* Hierbei ist der Umstand, daß die betrachteten Ereignisse den Charakter der Zufälligkeit besitzen sollen, nur insoweit von Bedeutung, als er dazu dient, für die jedesmal auf-

gestellten denkbaren Einzelfälle die Voraussetzung der gleichen Möglichkeit, d. h. der gleichen Berechtigung, einzuführen. Infolgedessen sind die Lehrsätze und Methoden der W.-R. auch noch bei solchen Aufgaben anwendbar, wo es sich überhaupt nur um Mengen von gleichberechtigten Dingen, ohne Rücksicht auf Zufälligkeit oder Gesetzmäßigkeit, handelt. Das tritt besonders deutlich bei den sogenannten geometrischen Wahrscheinlichkeiten hervor, von denen wir später einige Beispiele behandeln werden, denn bei diesen hat man in Wahrheit reine Häufigkeitsaufgaben vor sich, denen das Gewand der Zufälligkeit nur lose umgehängt worden ist. Ein solches Verhalten ist übrigens keine Besonderheit der W.-R., findet sich vielmehr auch in anderen Teilen der angewandten Mathematik wieder. So läuft z. B. der Inhalt der analytischen Mechanik, wie er in den Lehrbüchern gewöhnlich dargestellt wird, auf die Untersuchung gewisser Klassen von Differentialgleichungen hinaus. Ebenso ist die gewöhnliche Theorie der Linsensysteme in Wahrheit nur ein Teil der Liniengeometrie in optischem Gewande.

Der Inhalt der W.-R. besitzt hiernach als mathematische Methode eine selbständige Bedeutung, die ganz unabhängig davon ist, ob und wieweit die Sätze der W.-R. auf konkrete Vorgänge angewendet werden dürfen. Da ferner die Größe $\mathfrak{W}(E)$, die den Ausgangspunkt bildet, ein einfacher arithmetischer Begriff ist, so hat man, wie bereits früher angedeutet wurde, nicht zu befürchten, daß die an diesen Begriff geknüpften mathematischen Entwicklungen auf innere Widersprüche führen könnten. Zugleich ersieht man aber auch, daß in der Definition des Begriffes $\mathfrak{W}(E)$ nichts liegt, woraus sich — um es kurz auszudrücken — seine praktische Brauchbarkeit ohne weiteres als notwendige Folge ergäbe, daß vielmehr zu jener Definition noch etwas hinzukommen muß, sobald man zu Anwendungen der W.-R. übergehen will. Es entsteht daher die Frage, wie diese noch notwendige Ergänzung beschaffen sei.

Um die Vorstellung zu fixieren, denke man sich elf Urnen U_0, U_1, U_2, ... U_{10} mit je zehn Kugeln gegeben, ferner sei die Anzahl der weißen Kugeln jedesmal gleich der Indexnummer der betreffenden Urne, dann sind die $\mathfrak{W}\mathfrak{W}$. für einen weißen Zug bei den einzelnen Urnen der Reihe nach gleich den Zahlen 0.0, 0.1, 0.2, ... 1.0. Andererseits wird niemand Bedenken tragen, zuzugeben, daß die Stärke des Grundes für die Erwartung eines weißen Zuges mit der Nummer der Urne wachse, daß also dem größeren $\mathfrak{W}(E)$ der stärkere Erwartungsgrund entspreche, und umgekehrt. Hierdurch wird aber zunächst nur verständiich gemacht, wie man auf das Wahrscheinlichkeitsmaß $\mathfrak{W}(E)$ gekommen ist, dagegen ist ein Nachweis für die wirkliche Brauchbarkeit der Größen $\mathfrak{W}(E)$ damit noch nicht erbracht. Hierzu ist vielmehr etwas anderes erforderlich.

§ 10. Soll die Erwartungsschätzung $\mathfrak{W}(E)$ einen praktischen Wert besitzen, so muß sie sich in der Erfahrung bewähren. Eine solche Bewährung kann allerdings nicht durch einen einzelnen Versuch, sondern nur durch ganze Reihen von Versuchen bewirkt werden, denn wenn z. B. eine Urne neben einer Million schwarzer Kugeln nur eine einzige weiße enthält, so kann alle Stärke der Erwartung eines schwarzen Zuges nicht verhindern, daß bei einem einzelnen Versuche nun doch gerade die weiße Kugel zum Vorschein kommt. Das führt uns denn endlich zu derjenigen Voraussetzung, auf der alle ernsthaften Anwendungen der W.-R. beruhen, nämlich zu dem Satze von der Ausgleichung des Zufalls oder, genauer gesprochen, zu dem Satze von der *gleichmäßigen Erschöpfung der möglichen Fälle.*

Um das Wesen der genannten Voraussetzung deutlich zu machen, betrachten wir folgenden Fall. Ein bestimmter Versuch, bei dem im ganzen c gleichberechtigte und einander ausschließende Möglichkeiten vorhanden sind, werde unter konstant bleibenden Versuchsbedingungen unbegrenzt oft wiederholt, wie das ja z. B. beim Würfeln oder beim Ziehen aus einer Urne mit Zurücklegung der Kugel in weitem Umfange ausführbar ist. Ferner unterscheide man die einzelnen Möglichkeiten laufend durch die Nummern $1, 2, \ldots c$ und bezeichne mit $F(p)$ das Ereignis, das zu der durch die Nummer p angezeigten Möglichkeit gehört. Endlich sei $H(p, n)$ die rH., die für das Ereignis $F(p)$ beobachtet worden ist, wenn man sich auf die n ersten Versuche der ganzen Reihe beschränkt. Dann werden — das ist der Inhalt des in Rede stehenden Satzes — die Glieder der unbegrenzten Reihe

$$H(p, 1), \quad H(p, 2), \quad H(p, 3), \ldots$$

gegen den Grenzwert $1 : c$ konvergieren, so daß eine gleichmäßige Erschöpfung der einzelnen gleichberechtigten Möglichkeiten und damit auch für die beobachteten rHH. eine Ausgleichung der dem Einzelversuche anhaftenden Zufälligkeit eintritt. Infolgedessen wird ferner, wenn unter den c Möglichkeiten a für ein bestimmtes Ereignis E günstig sind, die beobachtete rH. dieses Ereignisses gegen den Wert $a : c$ oder $\mathfrak{W}(E)$ konvergieren. Weiter liegt in dem Satze, daß bei endlichen Versuchsreihen wenigstens eine angenäherte Ausgleichung des Zufalls erfolgen wird, wobei natürlich der Grad der Annäherung je nach den Umständen sehr verschieden ausfallen kann, da ja der Satz über die Art der Konvergenz keine bestimmte Aussage enthält. Endlich fließt noch aus dem Satze durch Umkehrung die Folgerung, daß Ereignisse, die mit gleicher rH. beobachtet worden sind, näherungsweise auch die gleiche Möglichkeit des Eintretens besitzen.

Fragt man jetzt nach den Gründen für die Berechtigung des Satzes, so läßt sich diese sicherlich nicht aus dem Begriffe der Zufälligkeit herleiten, die nach Voraussetzung den Ergebnissen der Einzel-

versuche anhaftet. Denn wenn die Einzelergebnisse dem Zufall unterworfen sind, so liegt darin ausgesprochen, daß zwischen ihnen keinerlei Zusammenhang besteht, daß also kein Ergebnis auf die folgenden einen Einfluß ausübt. Dann ist es aber schlechterdings unvorstellbar, wie dem Zufall die Tendenz innewohnen könne, sich selber auszugleichen. Ebensowenig kann man von der rein arithmetischen Definition der Größe $\mathfrak{W}(E)$ ausgehend durch irgend eine logische Deduktion den Satz erhärten, daß dieser Größe innerhalb wirklich ausgeführter Versuchsreihen nun auch mit Notwendigkeit eine reale Bedeutung im Sinne einer gleichmäßigen Erschöpfung der möglichen Fälle zukommen müsse. Der Satz, daß das, was gleich möglich ist, auch gleich oft eintreten werde, läßt sich nicht aus der bloßen Voraussetzung herleiten, daß gleich mögliche Fälle vorliegen.

Unter solchen Umständen bleibt nur übrig, die Begründung des Satzes auf die Aussagen der Erfahrung zu stützen und ihn als eine durch *Induktion* gewonnene Verallgemeinerung *beobachteter* Tatsachen anzusehen. Durch die mannigfachen Prüfungen, denen der Satz im Laufe der Zeit fortwährend unterworfen ist, hat seine Gültigkeit reichlich dasselbe Maß von Sicherheit erlangt, das wir zahlreichen anderen Erfahrungsgesetzen zuzusprechen gewohnt sind. Wo einmal, was ja möglich ist, ein zweifelloser Widerspruch zwischen der Vorschrift des Satzes und einer bestimmten Beobachtung auftritt, werden wir daraus nicht schließen, daß der Satz an sich falsch sei, sondern nur folgern, daß besondere, für gewöhnlich nicht eintretende Umstände als Ursache des Widerspruchs vorhanden sind. Es ist das dieselbe Maxime, der man auch anderwärts folgt: daraus, daß z. B. die Flugbahn des Bumerang den gewöhnlichen Fallformeln widerspricht, zieht man nicht den Schluß, daß diese Formeln an sich falsch seien.

In dem oben (§ 6) genannten Buche hat *J. von Kries* unter dem Namen „Prinzip der Spielräume“ eine logische Konstruktion entworfen, die geeignet ist, den Mechanismus deutlich zu machen, durch den die Ausgleichung des Zufalls zustande kommt. Da jedoch die Behandlung dieses Gegenstandes ein ziemlich weites Ausholen verlangen würde, so will ich hier nicht näher darauf eingehen, zumal jene Konstruktion den empirischen Charakter des Ausgleichungssatzes unangetastet läßt.

§ 11. Nach den vorstehenden Erörterungen nehmen wir die Betrachtung der Größen $\mathfrak{W}(E)$ wieder auf. Die erste Forderung, die bei der Berechnung dieser Größen zu erfüllen ist, besteht darin, daß die Reihe der gleichberechtigten Fälle in einwandfreier Weise aufgestellt werde. Wird z. B. mit einer aufgeworfenen Münze das Spiel „Bild oder Schrift“ ausgeführt, und fragt man nach der $\mathfrak{W}$., zweimal hintereinander Bild zu werfen, so kann man zunächst drei Fälle unterscheiden, nämlich: 1) zweimal Bild, 2) einmal Bild und einmal

Schrift, 3) zweimal Schrift. Bei dieser Einteilung würde man für die gesuchte W. den Wert 1 : 3 erhalten. Dabei ist aber außer acht gelassen, daß der zweite Fall auf zwei verschiedene Arten eintreten kann und daß man unter Berücksichtigung der zeitlichen Folge der Würfe die viergliedrige Reihe Bild—Bild, Bild—Schrift, Schrift—Bild und Schrift—Schrift erhält. Es ist nun nicht zweifelhaft, daß man von der zweiten Einteilung auszugehen und die gesuchte W. gleich 1 : 4 zu setzen hat, sobald man, wie üblich, die Annahme zugrunde legt, daß erstlich zwischen den aufeinander folgenden Würfen kein Zusammenhang bestehe, und daß zweitens bei dem einzelnen Wurfe die Fälle Bild und Schrift als gleich möglich zu gelten haben.

Die richtige Ansetzung der gleichberechtigten Fälle bietet mitunter erhebliche Schwierigkeiten dar, so daß der Versuch, die verlangten W.-Größen direkt zu berechnen, einfach scheitert. Als Erläuterung möge das gewöhnliche Würfelspiel dienen. Wenn zunächst der zum Spielen benutzte Würfel ein homogener und vollkommen regelmäßiger Sechsflächner ist, so wird man die sechs möglichen Würfe ohne Besinnen als gleichberechtigt nebeneinander stellen, denn einerseits läßt sich gegen die Gleichberechtigung kein stichhaltiger Grund vorbringen, andererseits kann man gegen jeden anderen, zur Berechnung der gesuchten WW. bestimmten Ansatz die gegebene Beschaffenheit des Würfels geltend machen. Es führt also hier das Prinzip des mangelnden Grundes sofort zur Lösung der gestellten Aufgabe.

Wenn der Würfel, wie es in Wirklichkeit stets der Fall sein wird, nur näherungsweise die soeben vorausgesetzte Beschaffenheit besitzt, so pflegt man wiederum die sechs möglichen Würfe ohne Bedenken als gleichberechtigt nebeneinander zu stellen, obgleich in diesem Falle das Prinzip des mangelnden Grundes tatsächlich versagt, da sich ja bestimmte Gründe gegen die Gleichberechtigung der Würfe geltend machen lassen. Man kann sich jedoch hierbei immer noch darauf berufen, daß die Gleichberechtigung der sechs Würfe wenigstens als Annäherung zugelassen werden dürfe, mit dem Vorbehalte, daß auch die aus einem solchen Ansatze berechneten Resultate ebenfalls nur als Annäherungen zu gelten haben.

Endlich denken wir uns den Fall, daß der zum Würfeln benutzte Körper nach Form und Struktur völlig unregelmäßig beschaffen sei, jedoch auf ebener Unterlage gerade sechs verschiedene stabile Ruhelagen annehmen könne, die die sechs möglichen Würfe liefern. Dann wird die Gleichberechtigung der sechs Würfe im allgemeinen so stark von der Wahrheit abweichen, daß sie nicht einmal mehr als rohe Annäherung zulässig ist. Der gleiche Fall würde auch noch eintreten, wenn der zum Würfeln benutzte Körper zwar einen merklich homogenen Sechsflächner mit rechteckigen Seiten darstellt, wenn aber die Kantenlängen stark ungleich sind, also z. B. sich wie 2 : 3 : 4

verhalten. Unter solchen Umständen ist man darauf angewiesen, die gesuchten $\mathfrak{W}$-Größen aus besondern ad hoc angestellten Beobachtungsreihen zu ermitteln; die Grundlagen dieser empirischen Bestimmungsweise werden wir später kennen lernen.

Die betrachteten drei Fälle können als Erläuterung des Satzes dienen, daß zu einem *brauchbaren* Wahrscheinlichkeitsansatze vor allem ein *bestimmtes objektives Wissen* nötig ist. Wo man nichts weiß, kann, wie wir es bereits früher ausgedrückt haben, auch die Kunst des Mathematikers nichts herausholen. Diese Bemerkung ist, soviel ich sehen kann, zuerst durch *J. von Kries* mit dem gebührenden Nachdruck hervorgehoben worden, während manche Schulbeispiele der Lehrbuchliteratur geradezu darauf hinauslaufen, bei dem Anfänger die Vorstellung zu erzeugen, daß die W.-R. ein Mittel sei, um die von dem Nichtwissen gelassenen Lücken auszufüllen.

§ 12. Bei der Berechnung einer $\mathfrak{W}(E)$ werden die einzelnen möglichen Fälle, ohne Rücksicht auf ihre besondere Beschaffenheit, lediglich darnach ausgezählt, ob sie für E günstig oder ungünstig sind. Infolgedessen ist es erlaubt, zur Veranschaulichung zufälliger Ereignisse einen bestimmten Vorgang als feststehendes Schema zugrunde zu legen. Hierzu wird von altersher mit Vorliebe das Ziehen von Kugeln aus einer Urne benutzt, das wir im folgenden ebenfalls gebrauchen wollen, obgleich mit Rücksicht auf die bequeme Ausführung die Anwendung besonders geformter Würfel mitunter den Vorzug verdienen würde. Man hat dabei die Füllung jedesmal so zu wählen, daß z. B. die $\mathfrak{W}$. eines weißen Zuges mit der $\mathfrak{W}$. des gerade betrachteten Ereignisses übereinstimmt. Um Wiederholungen zu vermeiden, wollen wir für die nächstfolgenden Abschnitte festsetzen, daß die Urne a weiße, b schwarze, zusammen also $c = a + b$ Kugeln enthalte. Ferner soll E einen weißen, F einen schwarzen Zug bedeuten.

Der herkömmlichen Festsetzung

$$\mathfrak{W}(E) = a : c, \qquad \mathfrak{W}(F) = b : c$$

gegenüber kann man fragen, weshalb gerade die rHH. als Wahrscheinlichkeitsmaß gewählt worden seien. Die Antwort darauf lautet, daß man ganz gut auch andere homogene Verbindungen der Zahlen a und b, z. B. die Quotienten $a : b$ und $b : a$, hätte einführen können, denn die Formeln und Lehrsätze der W.-R. hätten dadurch zwar ihr Aussehen, aber nicht ihr Wesen geändert. Im übrigen hat der Erfolg gelehrt, daß das Rechnen mit den rHH. im allgemeinen auf die bequemsten Formen führt. Ich beschränke mich deswegen hier darauf, eine Verbindung zu erwähnen, die später bei der Untersuchung der Kollektivgegenstände eine Rolle spielen wird. Handelt es sich um das Ereignis E, so kann man die Differenz $a - b$ als den Überschuß der günstigen Fälle über die ungünstigen bezeichnen. Dividiert man

durch c, so erhält man den *„relativen Überschuß“*. Bezeichnet man diesen mit u, so folgt aus

$$u = (a - b) : c, \qquad \mathfrak{W}(E) = a : c, \qquad 1 - \mathfrak{W}(E) = b : c$$

die Gleichung

$$u = 2\,\mathfrak{W}(E) - 1.$$

Die Größe u ist hiernach zwischen den Grenzen ± 1 enthalten.

Aus der Definition der $\mathfrak{W}$-Größen folgt noch unmittelbar der Satz, daß sich $\mathfrak{W}(E)$ und $\mathfrak{W}(\text{nicht } E)$ jedesmal zu Eins ergänzen.

§ 13. Die Größe $\mathfrak{W}(E)$ ist ihrer Bedeutung nach ein echter Bruch, der auch die Grenzwerte Null oder Eins annehmen kann, da ja die Fälle $a = 0$ oder $b = 0$ nicht ausgeschlossen sind. Die für $a = 0$ eintretende Gleichung $\mathfrak{W}(E) = 0$ besagt, daß das Ereignis E unmöglich ist, denn wenn die Urne keine weiße Kugel enthält, so ist ein weißer Zug unmöglich. Ebenso besagt die für $b = 0$ eintretende Gleichung $\mathfrak{W}(E) = 1$, daß das Ereignis E gewiß ist, denn wenn keine schwarze Kugel vorhanden ist, so muß ausnahmslos jeder Zug eine weiße Kugel liefern.

Hiernach treten die für gewöhnlich streng geschiedenen Begriffe Unmöglichkeit, Wahrscheinlichkeit und Gewißheit in der W.-R. als bloße Größenabstufungen eines und desselben Begriffes, nämlich der Größe $\mathfrak{W}(E)$ auf. Jedoch ist zu beachten, daß nicht immer, wenn der Gang der Rechnung auf die Gleichung $\mathfrak{W} = 0$ führt, daraus die logische Unmöglichkeit von E folgt. Man denke sich z. B. auf einer gegebenen geradlinigen Strecke zwei Punkte unabhängig voneinander nach Zufall gewählt, mit dem Hinzufügen, daß für alle Punkte der Strecke die gleiche Möglichkeit des Gewähltwerdens bestehen solle, dann zerfällt die betrachtete Gerade in drei Teilstrecken, die wir uns, soweit das möglich ist, zu einem Dreieck zusammengesetzt denken. Man erkennt nun sofort, daß das gleichseitige Dreieck nur in einem einzigen Falle unter unendlich vielen gleichberechtigten entstehen kann, daß also die $\mathfrak{W}$. für dieses Dreieck unendlich klein ausfällt und darum gleich Null zu setzen ist, wenn sie neben andere endliche $\mathfrak{W}$-Größen gestellt wird. Man hat demnach bei der Deutung der Gleichung $\mathfrak{W} = 0$ jedesmal darauf zu achten, ob sie aus dem absoluten Fehlen oder aber nur aus einer unendlich kleinen rH. des Ereignisses E entstanden ist.

§ 14. Bei der Berechnung der $\mathfrak{W}$-Größen wird fortwährend von zwei Lehrsätzen Gebrauch gemacht, die man passend als die *Summen-* und die *Produkt-Regel* bezeichnen kann. Um jedoch den Zusammenhang nachher nicht unterbrechen zu müssen, mögen der Herleitung jener Regeln die nachstehenden Festsetzungen vorausgeschickt werden.

Soll das gleichzeitige oder sukzessive Eintreten zweier Ereignisse E und E' ins Auge gefaßt werden, so spricht man von einem *aus* E

und E' zusammengesetzten Ereignisse, wofür wir kurz das Zeichen (EE') gebrauchen werden. So setzt sich z. B. ein zweimaliger weißer Zug aus einer Urne zusammen aus den beiden Ereignissen: E = weiß beim ersten Zuge und E' = weiß beim zweiten Zuge. In der gleichen Weise kann man von einem aus beliebig vielen Ereignissen $E, E', E'', \ldots$ zusammengesetzten Ereignisse $(EE'E''\ldots)$ sprechen.

Zwei Ereignisse E und E' heißen voneinander *unabhängig*, wenn das Eintreten von E das Eintreten von E' nicht beeinflußt und umgekehrt. Im entgegengesetzten Falle spricht man von einer *Abhängigkeit*, die zwischen E und E' besteht. Da man es bei den Anwendungen der W.-R. stets mit physischen Vorgängen zu tun hat, so ist die etwa auftretende Abhängigkeit immer nur einseitig, d. h. wenn E' von E abhängt, so hängt nicht auch umgekehrt E von E' ab. Das abhängige Ereignis ist darum auch immer das zeitlich folgende. So sind z. B. zwei Züge, die aus verschiedenen Urnen erfolgen, von einander unabhängig, desgleichen auch zwei Züge aus derselben Urne, falls die gezogene Kugel jedesmal zurückgelegt wird. Legt man dagegen die Kugeln nicht zurück, so ist der zweite Zug von dem ersten abhängig, denn der zweite Zug trifft auf das Farbenverhältnis $(a-1):b$ oder $a:(b-1)$, je nachdem der erste Zug weiß oder schwarz geliefert hat.

Die vorstehenden Unterscheidungen kann man natürlich auf Gruppen von beliebig vielen Ereignissen ausdehnen.

Dritte Vorlesung.

Allgemeine Lehrsätze.

§ 15. Bei der Herleitung der im letzten Paragraphen erwähnten Lehrsätze beginnen wir mit der Summenregel. Sie entsteht, wenn man die Gesamtheit der möglichen Fälle nach irgendwelchen Gesichtspunkten in Gruppen zerlegt und dem entsprechend auch die günstigen Fälle gruppenweise zusammenfaßt. Um diese Zusammenfassung an dem Urnenschema deutlich zu machen, denken wir uns, daß in der Urne U mit den Füllungszahlen a, b, c jede der c Kugeln außer durch ihre Farbe auch noch durch eine Nummer aus der Reihe der Zahlen $1, 2, \ldots n$ gekennzeichnet sei. Außerdem nehmen wir an, daß verschiedene Kugeln, ohne Rücksicht auf ihre Farbe, dieselbe Nummer tragen können, wie das ja, wenn c größer als n ist, von selber mit Notwendigkeit eintritt. Dann läßt sich zunächst die Gesamtheit der gleichmöglichen Züge in n Gruppen ordnen, indem man die Züge mit gleicher Nummer zusammenfaßt. Bedeutet ferner a_h

die Menge der weißen Kugeln mit der Nummer h, und bezeichnet man mit E_h das Ereignis E, wenn die gezogene weiße Kugel die Nummer h trägt, so besitzt $\mathfrak{W}(E_h)$ den Wert $a_h : c$. Andererseits ist die Gesamtheit der weißen Kugeln gleich

$$a = a_1 + a_2 + \cdots + a_n.$$

Daraus folgt, weil $\mathfrak{W}(E) = a : c$ ist, durch Division mit c die *Summenregel* in der Gestalt

$$\mathfrak{W}(E) = \mathfrak{W}(E_1) + \mathfrak{W}(E_2) + \cdots + \mathfrak{W}(E_n).$$

Hierin werden die Größen $\mathfrak{W}(E_h)$ als die *partialen* $\mathfrak{W}\mathfrak{W}$. bezeichnet, im Gegensatze zu der durch $\mathfrak{W}(E)$ gegebenen *totalen* $\mathfrak{W}$. des Ereignisses E.

Streift man die benutzte Einkleidung ab, so läßt sich der gefundene Satz auch folgendermaßen fassen. Für das Eintreten von E bestehen bei einer vorgelegten Aufgabe im ganzen n verschiedene und einander ausschließende Möglichkeiten, die wir mit M_1, M_2, ... bezeichnen; ferner werde durch E_h das Ereignis E bezeichnet, wenn sein Eintreten der Möglichkeit M_h entspringt; dann kommt jedem E_h eine bestimmte partiale $\mathfrak{W}(E_h)$ zu, und man erhält das gesuchte $\mathfrak{W}(E)$ oder die totale $\mathfrak{W}$. von E, indem man die sämtlichen partialen $\mathfrak{W}\mathfrak{W}$. summiert.

§ 16. Um die Produktregel abzuleiten stellen wir uns die Aufgabe, die $\mathfrak{W}$. eines zusammengesetzten Ereignisses (EE') durch die $\mathfrak{W}\mathfrak{W}$. der Bestandteile E und E' auszudrücken. Hierbei sind zwei Fälle zu unterscheiden, je nachdem zwischen E und E' eine Abhängigkeit besteht oder nicht.

Sind E und E' voneinander unabhängig, so setzen wir für E den Zug einer weißen Kugel aus einer Urne U mit den Füllungszahlen a, b, c, ebenso für E' den Zug einer weißen Kugel aus einer anderen Urne U' mit den Füllungszahlen a', b', c'. Man erhält dann die Gesamtheit der gleichmöglichen (günstigen und ungünstigen) Fälle, die bei den beiden Zügen aus U und U' in Betracht kommen, wenn man jeden Fall bei U mit jedem Fall bei U' kombiniert. Die Anzahl dieser Kombinationen ist cc'. Ferner erhält man die Gesamtheit der für (EE') günstigen Fälle, wenn man jede weiße Kugel in U mit jeder weißen Kugel in U' kombiniert. Die Menge dieser günstigen Kombinationen ist aa'. Daraus folgt für $\mathfrak{W}(EE')$ der Ausdruck $aa' : cc'$. Da nun andererseits die Werte von $\mathfrak{W}(E)$ und $\mathfrak{W}(E')$ durch die Quotienten $a : c$ und $a' : c'$ gegeben sind, so erhält man die gesuchte *Produktregel* in der einfachen Gestalt

$$\mathfrak{W}(EE') = \mathfrak{W}(E) \cdot \mathfrak{W}(E').$$

Ist ein dreigliedriges Ereignis $(EE'E'')$ mit den voneinander

unabhängigen Bestandteilen E, E', E'' gegeben, so zerlege man dieses zunächst in die beiden Bestandteile (EE') und E'', woraus

$$\mathfrak{W}(EE'E'') = \mathfrak{W}(EE') \cdot \mathfrak{W}(E'')$$

und weiter

$$\mathfrak{W}(EE'E'') = \mathfrak{W}(E) \cdot \mathfrak{W}(E') \cdot \mathfrak{W}(E'')$$

folgt. In der gleichen Weise erhält man durch wiederholte Anwendung des gefundenen Satzes für ein aus beliebig vielen und voneinander unabhängigen Gliedern zusammengesetztes Ereignis die Darstellung

$$\mathfrak{W}(EE'E''\cdots) = \mathfrak{W}(E) \cdot \mathfrak{W}(E') \cdot \mathfrak{W}(E'') \cdots.$$

Bedeuten E', E'', ... unabhängige Wiederholungen eines und desselben Ereignisses E, also z. B. wiederholte weiße Züge aus U unter Zurücklegung der Kugel, und bezeichnet man ferner das n-malige Eintreten von E kurz durch E^n, so wird

$$\mathfrak{W}(E^n) = \mathfrak{W}(E)^n.$$

Die letzte Formel macht deutlich, wie sich die $\mathfrak{W}$. für die fortgesetzte Wiederholung eines bestimmten Ereignisses E mit der wachsenden Anzahl der Wiederholungen beständig vermindert, da ja die Zahl $\mathfrak{W}(E)$ ein echter Bruch ist. Im Zusammenhange damit wächst die $\mathfrak{W}$. dafür, daß das Gegenteil von E wenigstens einmal eintreten wird. Bedeutet z. B. E das Werfen einer Nicht-Sechs bei dem gewöhnlichen Würfel, so ist $\mathfrak{W}(E)$ gleich $5:6$, und man erhält

$$\mathfrak{W}(E) = 0.83, \quad \mathfrak{W}(E^2) = 0.69, \quad \mathfrak{W}(E^3) = 0.58, \quad \mathfrak{W}(E^4) = 0.48, \text{ usw.}$$

Hiernach würde man, wie es häufig ausgedrückt wird, schon mit Vorteil die Wette Eins gegen Eins halten können, daß bei vier Würfen die Sechs wenigstens einmal auftritt.

§ 17. Wenn in dem Ereignis (EE') der Bestandteil E' von E abhängig ist, so versagt die oben benutzte Beweisführung und bedarf einer Abänderung. Zu dem Ende denken wir uns für den ersten Zug wieder wie oben die Urne U mit den Füllungszahlen a, b, c zugrunde gelegt, setzen aber für den zweiten Zug eine Urne U' mit c' Kugeln voraus, deren Farbe vorläufig, d. h. vor Ausführung des Zuges aus U, weder weiß noch schwarz ist. Ferner wollen wir uns vorstellen, daß in dem Augenblicke, wo der Zug aus U erfolgt ist, von den Kugeln in U' eine gewisse Menge die weiße Farbe annimmt, und zwar in der Anzahl a' oder a'', je nachdem der Zug aus U das Ergebnis E oder F geliefert hat. Dann kommen im ganzen cc' Zugpaare in Betracht, von denen aa' für (EE') günstig sind, da ja nach dem Eintreten von E der zweite Zug aus einer Urne U' mit den Füllungszahlen a', $c'-a'$, c' erfolgt. Demnach besitzt $\mathfrak{W}(EE')$ den Wert $aa':cc'$. Versteht man ferner unter $\mathfrak{W}(E')$ jetzt die $\mathfrak{W}$., die dem Ereignis E' *nach* dem Eintreten von E zukommt, so erhält man dafür

den Wert $a':c'$. Da endlich $\mathfrak{W}(E)$ gleich $a:c$ ist, so entsteht die Gleichung

$$\mathfrak{W}(EE') = \mathfrak{W}(E) \cdot \mathfrak{W}(E'),$$

die äußerlich dieselbe Gestalt besitzt, wie in dem Falle der gegenseitigen Unabhängigkeit der beiden Ereignisse E und E', nur daß jetzt dem Zeichen $\mathfrak{W}(E')$ eine andere Bedeutung als vorhin zukommt.

Die hier für U' vorausgesetzte Abhängigkeit der Farbenverteilung von dem Ausfall des ersten Zuges läßt sich unschwer verwirklichen, wenn man statt der einen Urne U' zwei Urnen $U(E)$ und $U(F)$ mit den Füllungszahlen a', b', c' und a'', b'', c'' benutzt und festsetzt, daß der zweite Zug aus $U(E)$ oder $U(F)$ erfolgen soll, je nachdem der erste Zug das Ergebnis E oder F geliefert hat.

Ist in dem zusammengesetzten Ereignis $(EE'E''\ldots)$ das Ereignis E' von E, ferner E'' von E und E' usw. abhängig, so erhält man durch wiederholte Anwendung des gefundenen Satzes die Darstellung

$$\mathfrak{W}(EE'E''\cdots) = \mathfrak{W}(E) \cdot \mathfrak{W}(E') \cdot \mathfrak{W}(E'') \cdots,$$

worin $\mathfrak{W}(E')$ die $\mathfrak{W}$. von E' nach dem Eintreten von E, $\mathfrak{W}(E'')$ die $\mathfrak{W}$. von E'' nach dem Eintreten von E *und* E', usw. bedeutet. Das Urnenschema erhält dann folgende Anordnung. Der erste Zug erfolgt aus einer einzigen Urne U. Für den zweiten Zug kommen zwei Urnen $U(E)$ und $U(F)$ in Betracht, je nachdem der erste Zug das Ergebnis E oder F geliefert hat. Für den dritten Zug sind vier Urnen $U(EE')$, $U(EF')$, $U(FE')$, $U(FF')$ anzusetzen, von denen je nach dem Ausfall des ersten Zugpaares immer nur eine Anwendung findet; in dieser Weise geht es fort, bis das letzte Glied des vorgelegten Ereignisses erschöpft ist. Man erkennt hieraus, daß sich bei abhängigen Ereignissen das Ergebnis der Zusammensetzung im allgemeinen viel verwickelter gestalten wird, als bei den unabhängigen.

§ 18. Da wir weiterhin die Summen- und die Produkt-Regel fortwährend anzuwenden haben werden, so wird es genügen, wenn wir hier zur Erläuterung eine einfache und häufig wiederkehrende Aufgabe behandeln.

Wir denken uns Reihen von je r Versuchen angestellt, von denen jeder das Ereignis E oder sein Gegenteil F liefern kann. Von diesen Reihen betrachten wir eine, die das Ereignis $(E^m F^n)$, d. h. in bestimmter Reihenfolge m-mal das Ergebnis E und n-mal das Ergebnis F, ergeben hat, wobei natürlich $m + n$ gleich r ist. Dann kann man nach der $\mathfrak{W}$. fragen, mit der jenes Ereignis zu erwarten war. Bei der Beantwortung der Frage wollen wir voraussetzen, daß die Versuche unabhängig voneinander seien, und daß ferner bei dem Versuche mit der Nummer h

$$\mathfrak{W}(E) = p_h, \qquad \mathfrak{W}(F) = 1 - p_h = q_h$$

sei.

Zur Lösung der Aufgabe denken wir uns das Ereignis $(E^m F^n)$ vollständig ausgeschrieben, d. h. wir schreiben die r-gliedrige Kombination hin, die den Buchstaben E m-mal und F n-mal enthält, und zwar in derjenigen Anordnung, in der E und F in dem betrachteten Ereignis aufeinander folgen sollen. Dann erhält man nach der Produktregel die gesuchte W., wenn für jedes E oder F, das an der h-ten Stelle steht, entsprechend p_h oder q_h geschrieben und die neue Kombination einfach als Produkt gelesen wird. Die gesuchte W. enthält also m p-Faktoren mit ebensovielen verschiedenen Nummern und ferner n q-Faktoren mit den n übrigen Nummern. Sind im besondern die p_h einander gleich und zwar gleich p, so fallen auch die q_h zusammen und sind gleich $1 - p$. Dadurch wird dann

$$\mathfrak{W}(E^m F^n) = p^m (1 - p)^n,$$

also unabhängig von der Reihenfolge der E und F.

Bildet man mit den beiden willkürlichen Größen u und v das r-gliedrige Produkt

$$P = (up_1 + vq_1)(up_2 + vq_2) \cdots (up_r + vq_r),$$

so entspricht in der Entwickelung von P jedem Ereignis von der Form $(E^m F^n)$ ein bestimmtes Glied, das den Faktor $u^m v^n$ enthält. Ebenso entspricht umgekehrt jedem Gliede, das den Faktor $u^m v^n$ enthält, ein bestimmtes Ereignis von der Form $(E^m F^n)$.

Bezeichnet man jetzt mit $w(E^m F^n)$ die W. dafür, daß bei r Versuchen ohne Rücksicht auf die Reihenfolge E und F überhaupt m-mal und n-mal vorkommen, so hat man nach der Summenregel alle Glieder in P zusammenzunehmen, die den Faktor $u^m v^n$ enthalten. Die Summe dieser Glieder gibt nach Abwerfung der Potenzen von u und v die gesuchte w-Größe. Hiernach wird

$$P = \sum u^m v^n w(E^m F^n),$$

wo die Summation nach m von 0 bis r und gleichzeitig nach n von r bis 0 läuft. Die hier gegebene Darstellung der gesuchten WW. als Entwickelungskoeffizienten einer erzeugenden Funktion ist ein Kunstgriff, der in der W.-R. häufig Anwendung findet.

Sind alle p_h gleich p und alle q_h gleich q, so wird

$$P = (up + vq)^r,$$

woraus

$$w(E^m F^n) = (m + n)!\, p^m q^n : m!\, n!$$

folgt. Die w-Größen sind dann also nichts anderes, als die Glieder der Entwickelung des Binoms $(p + q)^r$.

Die gefundenen Sätze lassen sich unschwer erweitern. Wenn in in der r-gliedrigen Versuchsreihe bei jedem Versuche s Ereignisse $E, E', E'', \ldots$ zu unterscheiden sind, und wenn ferner deren WW.

bei dem h-ten Versuche der Reihe die Werte $\mathfrak{W}(E)_h$, $\mathfrak{W}(E')_h$, $\mathfrak{W}(E'')_h, \ldots$ besitzen, so lassen sich die verschiedenen Fälle übersichtlich in der Entwickelung einer erzeugenden Funktion zusammenfassen, die sich als das Produkt von r Faktoren der Gestalt

$$u\mathfrak{W}(E)_h + u'\mathfrak{W}(E')_h + u''\mathfrak{W}(E'')_h + \cdots, \qquad (h = 1, 2, \ldots r)$$

darstellt. Sind im besondern die Größen $\mathfrak{W}(E)_h, \ldots$ innerhalb der Versuchsreihe konstant, so ergeben sich die Größen w aus der Entwickelung der Potenz

$$[\mathfrak{W}(E) + \mathfrak{W}(E') + \mathfrak{W}(E'') + \cdots]^r$$

nach dem polynomischen Lehrsatze.

§ 19. Zum Abschlusse der bisherigen Erörterungen haben wir nun noch die Frage nach der $\mathfrak{W}$. der sogenannten *Ursachen* zufälliger Ereignisse zu behandeln. Nach dem gewöhnlichen Sprachgebrauche ist unter der Ursache eines Ereignisses E der Inbegriff A aller derjenigen Umstände zu verstehen, von denen das Ereignis überhaupt abhängt. Dieser Inbegriff A läßt sich nun bei den Ereignissen, mit denen wir es hier zu tun haben, jedesmal in zwei Bestandteile B und C spalten, und zwar auf folgende Art. Zunächst können wir nämlich als Bestandteil B den Inbegriff derjenigen Umstände abtrennen, welche dem betrachteten Ereignis seine mathematische Wahrscheinlichkeit verleihen. So ist z. B. bei einem normalen Würfel der Bestandteil B dadurch bestimmt, daß der Würfel ein homogener Kubus sein soll, ebenso wird bei einer Urne, die mit gleichartigen Kugeln gefüllt ist, B durch die Füllungszahlen der verschieden gefärbten oder numerierten Kugeln umgrenzt. Die Gesamtheit C der Umstände, die in A nach Abscheidung des Bestandteils B übrig bleiben, umfaßt dann das, was in dem Ereignis E mit dem Charakter der Zufälligkeit behaftet ist, also z. B. bei dem Würfel den Verlauf des Wurfes und bei der Urne das Mischen und Ziehen der Kugeln.

In der W.-R. besteht nun der Gebrauch, dem Bestandteile B die Bezeichnung *Ursache des Ereignisses E* zu geben, wodurch das Wort Ursache, dem gewöhnlichen Sprachgebrauche gegenüber, offenbar eine erhebliche Einengung erfährt. Der Name ist imgrunde genommen recht unglücklich gewählt, indessen scheint es schwierig zu sein, einen besseren Ausdruck einzubürgern. Wenn es sich in der W.-R. immer nur um Versuche über zufällige Ereignisse handelte, so wäre das Wort „Versuchsbedingungen“ ganz passend; es kommen jedoch in der W.-R. oft genug auch Aufgaben vor, bei denen der Ausdruck Versuch ziemlich gezwungen klingt. Besser wäre noch der in der Fehlertheorie sich von selber darbietende Name „Bestimmungsmodus“ oder, wie *J. von Kries* vorgeschlagen hat, „Entstehungsmodus“. Dieser Sachlage gegenüber wollen wir es bei der herkömmlichen Bezeichnung

bewenden lassen und weiterhin unter „Ursache eines Ereignisses E“ den Inbegriff der Umstände verstehen, von denen die Größe $\mathfrak{W}(E)$ abhängt.

Wenn in der oben für ein Ereignis E zugrunde gelegten Beziehung $A = B + C$ der Bestandteil B bekannt ist, so ist damit auch $\mathfrak{W}(E)$ als gegeben anzusehen. Kennt man dagegen B garnicht oder nur unvollständig, so muß man entweder auf die Ermittelung von $\mathfrak{W}(E)$ von vornherein verzichten oder aber versuchen, aus Beobachtungen zu dem von A abhängenden Ereignis E eine Aussage über B und $\mathfrak{W}(E)$ zu gewinnen. Allerdings wird der zweite Weg wegen des in A enthaltenen zufälligen Bestandteils C im allgemeinen nicht auf notwendige, sondern nur auf mehr oder minder wahrscheinliche Aussagen führen.

§ 20. Die herkömmliche Lösung der betrachteten Aufgabe geht auf *Bayes* zurück und wird deshalb gewöhnlich kurz als das *Bayes*sche Prinzip bezeichnet. Die bisher übliche Formulierung dieses Prinzips rührt von *Laplace* her und soll nachstehend zugrunde gelegt werden.

Soll aus der Beobachtung eines Ereignisses E ein Schluß auf dessen unbekannte Ursache U gezogen werden, so ist diese Aufgabe gleichbedeutend mit der Frage nach dem Werte von $\mathfrak{W}(E)$, denn an dem mit U bezeichneten unbekannten Dinge interessiert hier nur der Umstand, daß U der Größe $\mathfrak{W}(E)$ einen bestimmten Zahlenwert erteilt.

Da die Ursache U unbekannt ist, so ist man vorläufig darauf angewiesen, über U eine Reihe von Annahmen zu machen, und dann die Berechtigung festzustellen, die jeder Annahme zuzusprechen ist. Wir wollen diese Annahmen mit $H_1, H_2, \ldots H_n$ bezeichnen und unter $\mathfrak{W}(EH_h)$ die $\mathfrak{W}$. verstehen, die dem betrachteten Ereignis E zukommen würde, wenn die Annahme H_h tatsächlich zutrifft. Die Reihe der H ist so anzusetzen, daß die einzelnen Annahmen einander ausschließen und in ihrer Gesamtheit die Möglichkeiten erschöpfen, die bei dem gerade vorgelegten Falle zu berücksichtigen sind. Ferner setzen wir fest, daß die Größen $\mathfrak{W}(EH)$ sämtlich voneinander verschieden sein sollen, da ja zwei Annahmen, die dasselbe $\mathfrak{W}(EH)$ liefern, für die vorliegende Betrachtung als nicht wesentlich verschieden, d. h. also als zusammenfallend anzusehen sind.

Die ziffernmäßige Abschätzung der relativen Berechtigungen denken wir uns in Gestalt einer Abstimmung vorgenommen: wenn von N gleichberechtigten und gültigen Stimmen, die überhaupt abgegeben werden können, N_h für die Annahme H_h sprechen, so soll die Berechtigung von H_h durch den Quotienten $N_h : N$ gemessen werden. Dieser Quotient ist eine rH. und weist alle Merkmale auf, die den $\mathfrak{W}$-Größen zukommen. Deshalb bezeichnen wir ihn durch das Symbol $\mathfrak{W}(H_h)$ und nennen ihn die „$\mathfrak{W}$. der Annahme H_h“. Die Summe aller $\mathfrak{W}(H)$ ist hiernach stets gleich Eins, da ja die Summe der N_h gleich N ist.

Das Ergebnis der Abstimmung muß im allgemeinen verschieden ausfallen, je nachdem sie vor oder nach dem Bekanntwerden des Ereignisses E vorgenommen wird. Weiß man z. B. vorläufig nur, daß eine Urne 100 Kugeln enthält, während die Farbenverteilung zunächst unbekannt ist, so hat man vorerst keinen Grund, die Annahme „2 weiße und 98 schwarze Kugeln“ auszuschließen, und wird ihr demgemäß vorläufig ein von Null verschiedenes $\mathfrak{W}(H)$ zusprechen. Man wird sogar in dem betrachteten Falle keinen Anstand nehmen, jeder der möglichen Annahmen vorläufig das gleiche $\mathfrak{W}(H)$, nämlich den Wert $1:101$ zu erteilen. Besteht nun aber das beobachtete Ereignis E darin, daß ohne Zurücklegen der Kugel 3 weiße Züge erfolgt sind, so wird dadurch die betrachtete Annahme nachträglich ausgeschlossen, d. h. man erhält für $\mathfrak{W}(H)$ nunmehr den Wert Null. Demgemäß werden wir die vor und nach dem Bekanntwerden von E erfolgenden Abschätzungen durch die Symbole $\mathfrak{W}(H)_v$ und $\mathfrak{W}(H)_n$ unterscheiden und sie als die *vorläufigen* und die *nachträglichen* $\mathfrak{W}\mathfrak{W}$. der Annahmen H bezeichnen.

§ 21. Es handelt sich jetzt darum, die Beziehungen zwischen den Größen $\mathfrak{W}(EH)$, $\mathfrak{W}(H)_v$ und $\mathfrak{W}(H)_n$ aufzufinden, wobei wir wieder das Urnenschema zugrunde legen. Gegeben sind m Urnen, deren jede die gleiche Anzahl von Kugeln, nämlich c, enthalten soll. Der Fall, daß alle Urnen die gleiche Farbenverteilung besitzen, wird ausgeschlossen, dagegen wird zugelassen, daß zwei oder mehr Urnen in der Farbenverteilung übereinstimmen. Geht man die Urnenreihe Urne für Urne durch, so soll die Anzahl der weißen Kugeln im ganzen die n verschiedenen Werte $a_1, a_2, \ldots a_n$ aufweisen, so daß n kleiner als m sein kann, während der Fall $n = 1$ ausgeschlossen ist. Danach teilen wir die Urnenreihe in n Gruppen $H_1, \ldots H_n$, indem wir der Gruppe H_h alle Urnen zuweisen, die a_h weiße Kugeln enthalten. Hierbei komme der Gruppe H_h die Urnenmenge m_h zu, woraus für die Summe aller m_h der Wert m folgt. Das beobachtete Ereignis ist ein weißer Zug: gesucht wird die $\mathfrak{W}$. der Annahme, daß die gezogene Kugel aus einer Urne der Gruppe H_h stamme. Die Gruppen H vertreten hierbei offenbar die vorhin aufgestellten einzelnen Annahmen über die Ursache U.

Bei der Lösung der Aufgabe wollen wir die Voraussetzung zugrunde legen, daß die Urnen vor dem Bekanntwerden der gezogenen Farbe gleichberechtigt nebeneinander stehen, daß also für alle Urnen der gleiche Grad von Ungewißheit hinsichtlich der erfolgten Benutzung bestehe. Dann sind von den m vorhandenen Urnen vorläufig m_h für die Annahme H_h günstig, d. h. es ist

$$\mathfrak{W}(H_h)_v = m_h : m \tag{1}$$

zu setzen. Ferner läßt sich, so lange man die gezogene Farbe noch

nicht kennt, noch folgendes aussagen: es ist zwar gewiß, daß eine Kugel gezogen wurde, dagegen ist es vorläufig ungewiß, welche von den mc Kugeln, die im ganzen vorhanden sind, der Zug getroffen hat; außerdem ist diese Ungewißheit für alle Kugeln die gleiche, da ja jede Urne gleichviel Kugeln enthält, und da für alle Urnen vorläufig die gleiche Ungewißheit ihrer Benutzung besteht. Durch das Bekanntwerden des Zuges wird nun die für die mc Kugeln bestehende Ungewißheit bezüglich der Farbe aufgehoben, sie bleibt aber für die weißen Kugeln in derselben Weise wie vorher bestehen, denn die bloße Tatsache des weißen Zuges bietet, wenn man sich innerhalb jeder Urne die weißen Kugeln laufend numeriert denkt, keinerlei Anhalt für die Aussage, daß z. B. eher die zweite Kugel aus der dritten Urne, als die fünfte Kugel aus der vierten Urne gezogen worden sei. Unter solchen Umständen sprechen die $m_h a_h$ weißen Kugeln, die zusammen in den Urnen der Gruppe H_h enthalten sind, für die Gruppe oder die Annahme H_h, d. h. es ist

$$\mathfrak{W}(H_h)_n \text{ proportional } m_h a_h. \tag{2}$$

Die $\mathfrak{W}$. des weißen Zuges ist, wenn man eine Urne der Gruppe H_h benutzt, durch den Quotienten $a_h : c$ gegeben, so daß mit der eingeführten Bezeichnung

$$\mathfrak{W}(EH_h) = a_h : c \tag{3}$$

wird. Demnach nimmt (2) unter der Berücksichtigung von (1) und (3) die Gestalt

$$\mathfrak{W}(H_h)_n \text{ proportional } [\mathfrak{W}(EH_h)\,\mathfrak{W}(H_h)_v]$$

an, wofür wir, da die Summe der $\mathfrak{W}(H)_n$ gleich Eins ist, auch

$$\mathfrak{W}(H_h)_n = [\mathfrak{W}(EH_h)\,\mathfrak{W}(H_h)_v] : P, \tag{4}$$

$$P = \sum_h \mathfrak{W}(EH_h)\,\mathfrak{W}(H_h)_v \tag{5}$$

schreiben können. Damit ist die gesuchte Beziehung gefunden. Sie enthält die Abgrenzung der relativen Berechtigung, die auf Grund der Beobachtung von E nachträglich jeder Annahme H zuzusprechen ist.

§ 22. Wenn von zwei Ereignissen E und E', die auf derselben Ursache U beruhen, das erste beobachtet ist, das zweite aber noch nicht, so gestattet der Schluß von E auf U einen weiteren Schluß auf $\mathfrak{W}(E')$. Bedeuten $H_1, \ldots H_n$ wieder die Annahmen, die über U zu machen sind, so hat man für das Eintreten von E' n Möglichkeiten zu unterscheiden, je nachdem nämlich die Annahme H_1 oder H_2 oder H_3 usw. zutrifft. Das Eintreten von E' nach der h-ten Möglichkeit erscheint hierbei als ein zusammengesetztes Ereignis, dessen erster Bestandteil durch das Zutreffen der Annahme H_h gebildet wird, während der zweite Bestandteil davon herrührt, daß das Ereignis E'

aus der Ursache H_h entspringt. Die h-te Art des Eintretens liefert deshalb nach der Produktregel die partiale $\mathfrak{W}$.

$$\mathfrak{W}(H_h)_n \, \mathfrak{W}(E'H_h),$$

wo der erste Faktor die aus dem beobachteten Ereignis E erschlossene $\mathfrak{W}$. der Annahme H_h bedeutet, wogegen der zweite Faktor die $\mathfrak{W}$. angibt, die dem Ereignis E' zukommt, wenn die Annahme H_h tatsächlich zutrifft. Daraus folgt für die totale $\mathfrak{W}$. von E' der Ausdruck

$$\mathfrak{W}(E') = \sum_h \mathfrak{W}(H_h)_n \, \mathfrak{W}(E'H_h). \tag{6}$$

Die Größe $\mathfrak{W}(E')$ erscheint hiernach als ein Mittelwert aus den $\mathfrak{W}(E'H)$, da ja die Summe der $\mathfrak{W}(H)_n$ den Betrag Eins besitzt.

§ 23. Wenn in den Formeln (4), (5) und (6) die Zahl der aufgestellten Annahmen unendlich groß ausfällt, wie das bei den Anwendungen häufig vorkommt, so ist es zweckmäßig, jene Gleichungen etwas umzugestalten. Wir setzen die Reihe der Annahmen zunächst als endlich voraus und bezeichnen mit n die Anzahl der Glieder dieser Reihe, mit h die Nummer eines Gliedes und mit x den Quotienten $h:n$. Dann verläuft die Veränderliche x zwischen den Grenzen 0 und 1 sprungweise mit dem konstanten Inkrement $1:n$. Mit Rücksicht darauf bezeichnen wir die zur Nummer h oder zu dem „Index“ x gehörige Annahme H_h weiterhin mit K_x und führen die unmittelbar verständlichen Zeichen

$$\mathfrak{W}(K_x)_v, \quad \mathfrak{W}(K_x)_n, \quad \mathfrak{W}(EK_x), \quad \mathfrak{W}(E'K_x) \tag{7}$$

ein. Wenn nun n über alle Grenzen wächst, so verwandelt sich x in eine stetige Veränderliche, deren konstantes Inkrement $1:n$ entsprechend die Bezeichnung dx erhält. Von den vier Größen (7) bleiben hierbei die letzten im allgemeinen endlich, während die beiden ersten im allgemeinen unendlich klein werden, weil ja ihre über alle Werte von x erstreckten Summen den Wert Eins besitzen. Demgemäß ersetzen wir die vier von x abhängenden Symbole (7) durch die neuen Zeichen

$$V(x)\,dx, \quad N(x)\,dx, \quad F(x), \quad G(x),$$

wobei

$$\sum \mathfrak{W}(K_x)_v = \int_0^1 V(x)\,dx = 1,$$

$$\sum \mathfrak{W}(K_x)_n = \int_0^1 N(x)\,dx = 1$$

wird. Da ferner aus (5) der Ausdruck

$$P = \sum \mathfrak{W}(EK_x)\,\mathfrak{W}(K_x)_v = \int_0^1 F(x)\,V(x)\,dx$$

entsteht, so nimmt (4) und (6) die Gestalt

$$P\mathfrak{W}(K_x)_n = PN(x)\,dx = F(x)\,V(x)\,dx, \tag{8}$$

$$P\mathfrak{W}(E') = \Sigma P\mathfrak{W}(K_x)_n \mathfrak{W}(E'K_x) = \int_0^1 F(x)\,G(x)\,V(x)\,dx \tag{9}$$

an.

Will man es durchweg mit endlichen Größen zu tun haben, so kann dies auf folgende Weise erreicht werden. Man bezeichne mit $\mathfrak{W}(X)_v$ und $\mathfrak{W}(X)_n$ die vorläufige und die nachträgliche $\mathfrak{W}$., daß für die gemeinsame Ursache von E und E' irgend eine der Annahmen K_x zutrifft, deren Index x zwischen o und X liegt, dann wird nach der Summenregel

$$\mathfrak{W}(X)_v = \int_0^X V(x)\,dx, \tag{10}$$

$$P\mathfrak{W}(X)_n = P\int_0^X N(x)\,dx = \int_0^X F(x)\,V(x)\,dx. \tag{11}$$

Der Index x hat in den vorstehenden Formeln offenbar nur den Zweck, die unendliche Mannigfaltigkeit der aufzustellenden Annahmen in eine feste Ordnung zu bringen und zugleich jedes Glied der Mannigfaltigkeit unzweideutig zu bezeichnen. Dieser Zweck läßt sich aber auf unendlich viele Arten erfüllen. Es sei y eine gewisse Funktion von x, die stetig wachsend das Intervall von o bis 1 durchläuft, sobald dies mit x der Fall ist. Dann gehört zu jedem x ein und nur ein bestimmtes y, und umgekehrt; ebenso gehört in der vorgelegten Mannigfaltigkeit der Annahmen zu jeder Annahme ein bestimmter Wert des neuen Index y, und umgekehrt. Wir wollen für die auf den neuen Index bezogenen Stücke die früheren Buchstaben, aber unter Anfügung eines Akzents, benutzen, so daß den Stücken

$$K_x,\quad V(x),\quad N(x),\quad F(x),\quad G(x),\quad \mathfrak{W}(X),$$

die neuen Zeichen

$$K'_y,\quad V'(y),\quad N'(y),\quad F'(y),\quad G'(y),\quad \mathfrak{W}'(Y)$$

entsprechen. Dann ist zunächst, wenn x und y ein simultanes Wertepaar der Indizes bedeuten, die Annahme K_x identisch mit K'_y, ferner wird $F(x)$ gleich $F'(y)$ und $G(x)$ gleich $G'(y)$. Sind auch X und Y simultane Indexwerte, so wird weiter

$$\mathfrak{W}(X)_v = \mathfrak{W}'(Y)_v,$$

also auch

$$\int_0^X V(x)\,dx = \int_0^Y V'(y)\,dy,$$

woraus durch Differentiation

$$V(X)\,dX = V'(Y)\,dY \tag{12}$$

folgt.

Als Beispiel denken wir uns den Fall, daß man bei einem Würfel von augenfällig unregelmäßiger Gestalt die 𝔚. für das Werfen der Eins empirisch dadurch bestimmen will, daß man eine Reihe von Würfen tatsächlich ausführt und die Häufigkeit von Eins und Nicht-Eins abzählt. Man kann dann die aufzustellenden Annahmen zunächst nach dem Werte x ordnen, den sie der 𝔚. für das einmalige Werfen der Eins erteilen. Dies ist der Ansatz, den man gewöhnlich wählt, und zwar aus dem einfachen Grunde, weil er der nächstliegende und bequemste ist. Ebensogut könnte man aber als ordnende Größe auch z. B. die 𝔚. für das zweimalige Werfen der Eins ansetzen, so daß der neue Index y gleich x^2 wird. Dieser Umstand ist, wie wir nachher sehen werden, für die richtige Beurteilung der gefundenen Formeln von Bedeutung.

§ 24. Die letzten Abschnitte enthalten die Formulierung des sogenannten *Bayes*schen Prinzips über die 𝔚. von Ursachen. Sie bieten einen passenden Anlaß, den bisher nicht berührten Unterschied zwischen der sogenannten *objektiven* und *subjektiven* Wahrscheinlichkeit zu erörtern. Bedeutet E ein Ereignis von derjenigen Beschaffenheit, welche in den bisherigen Betrachtungen fortwährend vorausgesetzt worden ist, so kommt E eine bestimmte 𝔚. zu, deren Wert ganz unabhängig davon ist, ob wir über die Ursache von E etwas wissen oder nicht. Das einfachste Beispiel hierfür ist der weiße oder schwarze Zug aus einer Urne, denn die 𝔚𝔚. dieser Züge hängen, gleichartige Kugeln vorausgesetzt, nur von den Füllungszahlen ab. Eine solche 𝔚. wird als die *objektive* Wahrscheinlichkeit oder auch als die *Chance* des Ereignisses E bezeichnet. Ist die Ursache von E vollständig bekannt, so kommt die Berechnung der Chance auf eine rein mathematische Aufgabe hinaus. Dieser Fall tritt, wenn nicht mit aller Strenge, so doch mit großer Annäherung im allgemeinen bei solchen Glücksspielen auf, deren Einrichtung und Anordnung vollständig in unserer Hand liegt. Ist uns dagegen die Ursache nicht vollständig zugänglich, so wird sich diese Lücke in unserem Wissen bei der Berechnung von $\mathfrak{W}(E)$ irgendwie geltend machen. Ein derart hergeleiteter Wert heißt dann *subjektive* Wahrscheinlichkeit und wird im allgemeinen von dem objektiven Werte verschieden sein.

Zur Erläuterung betrachten wir die in den Formeln (4) bis (9) auftretenden Bestandteile. Die Größen $\mathfrak{W}(EH)$, $\mathfrak{W}(E'H)$, $F(x)$, $G(x)$ sind offenbar objektiver Natur, denn sie bedeuten 𝔚-Größen, die aus *gegebenen* Ursachen berechnet werden. Dagegen ist $\mathfrak{W}(E')$ subjektiver Natur, denn diese Größe ist ein mit den Gewichten $\mathfrak{W}(H)_n$ aus den $\mathfrak{W}(E'H)$ berechneter Mittelwert, und die benutzten Gewichte enthalten subjektive, aus zwei verschiedenen Quellen entspringende Bestandteile. Erstens nämlich hängen die Zahlenwerte der $\mathfrak{W}(H)_n$ von dem Ausfall des Versuches ab, der zur Ermittelung der Ursache angestellt wurde

und das betrachtete Ereignis E lieferte. Wiederholt man den Versuch, so kann das Ergebnis anders ausfallen und damit eine Änderung aller weiteren, mit E zusammenhängenden Größen bewirken. Das individuelle Wissen, das dem Beobachter durch den angestellten Versuch vermittelt werden soll, ist also im allgemeinen unvollständig, und diese Unvollständigkeit bleibt in der Aussage bestehen, die der Beobachter über die Ursache von E bildet.

Der zweite subjektive Bestandteil entspringt aus der Ansetzung der Größen $\mathfrak{W}_v$. In der Regel wird dabei so verfahren, daß man erst die Reihe der Annahmen H aufstellt und dann sagt, daß die $\mathfrak{W}_v$ als einander gleich anzusetzen seien, weil man über die Ursache des beobachteten Ereignisses nichts wisse. Nun ist aber der Fall völliger Unwissenheit garnicht als der eigentlich normale anzusehen; oft genug ist man imstande, auf Grund früherer Erfahrung diese oder jene Aussage über die Ursache zu machen, und es dürfte nicht zweifelhaft sein, daß ein Schluß, der sich lediglich auf das „privilegium ignorantiae" stützt, weniger Vertrauen verdient, als ein anderer, bei dem ein bestimmtes, wenn auch vielleicht nur geringes, Wissen verwertet worden ist. Meistens ist es auch nur die Schwierigkeit, das etwa vorhandene Wissen ohne Hinzunahme willkürlicher Voraussetzungen zu verwerten, die den Anlaß gibt, daß man sich mit der bequemsten Festsetzung, nämlich der Einführung gleicher $\mathfrak{W}_v$ begnügt.

Des weiteren ist zu beachten, daß im Falle einer unendlichen Reihe von Annahmen die Auswahl des ordnenden Index x garnicht im voraus vorgeschrieben werden kann. Greifen wir auf das oben erwähnte Beispiel eines unregelmäßig geformten Würfels zurück, so ist es allerdings naheliegend, daß man als Index die $\mathfrak{W}$. für den einmaligen Wurf der Eins wählt. Mit demselben Rechte kann man aber auch z. B. die $\mathfrak{W}$. für das zweimalige Werfen der Eins, d. h. also die Größe $y = x^2$, als Index wählen. Wenn man ferner, unter Berufung auf unsere völlige Unkenntnis der Ursache, die $\mathfrak{W}$-Größen, d. h. also die Funktion $V(x)$, konstant setzt, so kann man mit demselben Rechte auch die Funktion $V'(y)$ konstant setzen, und man erkennt leicht, daß diese beiden Ansätze in dem weiteren Verlaufe der Rechnung zu verschiedenen Ergebnissen führen.

§ 25. Wie man sieht, liegt der schwache Punkt bei der Ermittlung der Größen $\mathfrak{W}_n$, $N(x)$, $\mathfrak{W}(E')$ weniger in dem Charakter der Zufälligkeit, der dem Ergebnis des angestellten Versuchs anhaftet, als vielmehr in der ausgeprägt subjektiven Beschaffenheit der $\mathfrak{W}_v$-Größen. Man kann nun fragen, ob sich dieser Mangel, wenn auch nicht beseitigen, so doch mildern lasse. Das ist in der Tat möglich, wie an späterer Stelle in der XIX. Vorlesung bei Behandlung der sogenannten *Bayes*schen Formel gezeigt werden wird.

Das sogenannte *Bayes*sche Prinzip, das in Wahrheit gar kein

Prinzip, sondern ein arithmetischer Lehrsatz ist, hat für die erkenntnistheoretische Erörterung der Grundlagen der W.-R. öfters einen Stein des Anstoßes gebildet, was allerdings leicht eintreten kann, wenn man hinter dem Worte „mathematische Wahrscheinlichkeit" mehr sucht, als den eingebürgerten Kunstausdruck für einen rein arithmetischen Begriff, nämlich für die bei gewissen Abzählungen auftretenden „relativen Häufigkeiten". Im übrigen erscheint das genannte Prinzip, wie sich später bei Erörterung der *Bayes*schen Formel zeigen wird, als ein durchaus entbehrlicher Bestandteil der W.-R., wenigstens wie sich die Dinge bisher entwickelt haben. Gleichwohl wird man es in einer Darstellung der W.-R. vorläufig nicht unterdrücken dürfen, und zwar der historischen Vollständigkeit wegen, denn es hat wenigstens einmal bei einer bestimmten Gelegenheit, nämlich bei der ersten *Gauß*schen Begründung der Methode der kleinsten Quadrate, als heuristisches Hilfsmittel einen nützlichen Dienst geleistet.[1])

§ 26. Die bisherigen Entwicklungen enthalten alles, was seither an allgemeinen Sätzen in den Anwendungen der W.-R. gebraucht worden ist. Rückblickend kann man sagen, daß diese Grundlagen einfach genug sind; die Schwierigkeiten, die sich bei besonderen Aufgaben einstellen können, haben, wie das auch anderwärts stattfindet, ihren Grund wesentlich in den besonderen Bedingungen eben dieser Aufgaben.

Bei der Beweisführung haben wir in ausgedehntem Maße von dem Urnenschema Gebrauch gemacht. Daraus folgt, daß die gefundenen Sätze unmittelbar nur bei solchen Aufgaben Anwendung finden können, die sich auf das genannte Schema reduzieren lassen. Der wesentliche Punkt liegt hierbei weniger in der Zufälligkeit der betrachteten Ereignisse, als vielmehr darin, daß eine $\mathfrak{W}$-Größe nichts anderes ist, als die rH. der günstigen Glieder in einer Reihe von gleichberechtigten, sich gegenseitig ausschließenden und voneinander unabhängigen Fällen. Die Ansetzung solcher Reihen ist nun keineswegs bei allen den Vorkommnissen möglich, bei denen wir von Zufall sprechen. Deshalb ist denn auch der Anwendungsbereich der W.-R. erheblich enger, als das Gebiet der Dinge, die uns mit Zufälligkeit behaftet zu sein scheinen. Dieser Umstand ist nicht immer genügend beachtet worden, und

1) Die vorstehende, an *Laplace* anknüpfende Darstellung der Grundbegriffe und Hauptlehrsätze der W.-R. ist mit gewissen Einschränkungen behaftet, die für die Anwendungen der W.-R. allerdings ohne Belang sind, und im übrigen dem Anfänger das Verständnis erleichtern. Diese Einschränkungen lassen sich jedoch aufheben, wenn man die W.-R. von Anfang an nur als eine „Häufigkeitsrechnung" behandelt und zugleich den Begriff der „relativen" Wahrscheinlichkeit einführt, wie das *Hausdorff* in seinem Aufsatze *„Beiträge zur Wahrscheinlichkeitsrechnung"* (Berichte der math.-phys. Klasse der k. Sächs. Ges. d. Wiss. 1901) getan hat. Das *Bayes*sche Prinzip erscheint dann erst recht als ein einfaches Korollar aus rein arithmetischen Lehrsätzen.

daraus sind dann Untersuchungen entsprungen, die mehr den Charakter einer unter Umständen höchst interessanten Rechenübung, als den einer ernsthaften Anwendung der W.-R. besitzen. In dem mehrfach genannten *von Kries*schen Buche finden sich an verschiedenen Stellen recht lehrreiche kritische Bemerkungen über solche mißbräuchliche Wahrscheinlichkeitsansätze. Allerdings läßt sich ja die Häufigkeitsrechnung überall da anwenden, wo es etwas zu zählen gibt, man muß sich aber hüten, die erlangten Zahlen als W.-Größen zu deuten, wenn die notwendigen Vorbedingungen dafür nicht erfüllt sind.

Vierte Vorlesung.

Die Transzendente $\Pi(x)$.

§ 27. Die Einfachheit der in den letzten Abschnitten entwickelten grundlegenden Sätze bringt es mit sich, daß zahlreiche Aufgaben der W.-R. mit elementaren Hilfsmitteln gelöst werden können. Es kommen indessen auch wichtige Aufgaben vor, deren bündige Behandlung ein wirksameres Werkzeug verlangt. Das, was hierbei nötig ist, findet sich in der Hauptsache in den besseren Lehrbüchern der Integralrechnung fertig vor oder läßt sich unschwer daraus herleiten. Ich halte es jedoch für zweckmäßig, das, was wir später an besonderen Lehrsätzen und Formeln der Integralrechnung brauchen, übersichtlich in den nächsten Abschnitten zusammenzustellen, um darauf unmittelbar Bezug nehmen zu können.

Zur Bezeichnung der gewöhnlich in der Gestalt e^x geschriebenen Exponentialfunktion werde das für den Druck vorteilhaftere Zeichen $\exp x$ verwendet und die Funktion $\Pi(x)$ durch die Gleichung

$$\Pi(x) = \int_0^\infty dt\, t^x \exp(-t) \tag{1}$$

definiert. Das Integral besitzt eine Bedeutung, so lange der Parameter x oberhalb -1 liegt. Bie Bedeutung von $\Pi(x)$ für Argumente unterhalb -1 wird sich weiterhin aus einer anderen Darstellung ergeben. Für $x = 0$ läßt sich die Integration unmittelbar ausführen und liefert

$$\Pi(0) = 1. \tag{2}$$

Wird die Gleichung

$$d(t^x \exp(-t)) : dt = -t^x \exp(-t) + x t^{x-1} \exp(-t)$$

links und rechts nach t zwischen den Grenzen 0 und ∞ integriert, wobei x als positiv zu denken ist, so erhält man

$$\Pi(x) = x\,\Pi(x-1). \tag{3}$$

Bedeutet n eine ganze positive Zahl, und setzt man in der vorstehenden Gleichung für x die Werte $1, 2, \ldots n$ ein, so gibt die Multiplikation der so entstehenden Gleichungen wegen (2)

$$\Pi(n) = 1 \cdot 2 \cdot 3 \cdots n = n!. \tag{4}$$

Mit Rücksicht hierauf wird $\Pi(x)$ auch als die allgemeine analytische Fakultät bezeichnet.

Bedeutet a eine positive Konstante, und wird in (1) au für t gesetzt, so erhält man

$$\int_0^\infty du\, u^x \exp(-au) = \Pi(x) : a^{x+1}. \tag{5}$$

Differentiiert man ferner (1) n-mal nach x, so wird

$$d^n \Pi(x) : dx^n = \int_0^\infty dt\, t^x \exp(-t)(\log t)^n. \tag{6}$$

Um hieraus die logarithmische Ableitung von $\Pi(x)$ zu erhalten, denken wir uns die leicht zu beweisende Gleichung

$$\int_0^\infty ds \exp(-st) = 1 : t, \quad (t > 0)$$

nach t zwischen den Grenzen 1 und t integriert, woraus zunächst

$$\log t = \int_0^\infty \frac{ds}{s}(\exp(-s) - \exp(-st))$$

folgt. Setzt man diesen Ausdruck von $\log t$ in (6) für den Fall $n = 1$ ein, so wird

$$\begin{aligned} d\Pi(x) : dx &= \int\!\!\int \frac{ds\, dt}{s} t^x \exp(-t)(\exp(-s) - \exp(-st)) \\ &= \int \frac{ds}{s} \exp(-s) \int dt\, t^x \exp(-t) - \int \frac{ds}{s} \int dt\, t^x \exp(-(1+s)t), \end{aligned}$$

woraus mit Rücksicht auf (1) und (5) für die logarithmische Ableitung $\Psi(x)$ die Darstellung

$$\Psi(x) = \int_0^\infty \frac{ds}{s}(\exp(-s) - (1+s)^{-x-1}) \tag{7}$$

folgt. Zieht man hiervon die für $x = 0$ entstehende Gleichung ab, so ergibt sich

$$\Psi(x) - \Psi(0) = \int_0^\infty \frac{ds}{s}((1+s)^{-1} - (1+s)^{-x-1}), \tag{8}$$

woraus mit der Substitution $1 + s = 1 : u$ die Gleichung

$$\Psi(x) - \Psi(0) = \int_0^1 du \frac{1-u^x}{1-u} \tag{9}$$

fließt. Entwickelt man hierin den Nenner nach steigenden Potenzen

von u, so läßt sich die Integration gliederweise unmittelbar ausführen, und man erhält

$$\Psi(x) - \Psi(0) = \sum_k \left(\frac{1}{k} - \frac{1}{k+x}\right), \quad (k = 1, 2, \ldots), \tag{10}$$

woraus durch Differentation für die Ableitung $\Psi'(x)$ die Reihe

$$\Psi'(x) = \sum_k (k+x)^{-2} \tag{11}$$

folgt. Andererseits liefert die Integration von (10) nach x zwischen den Grenzen 0 und x die Beziehung

$$\log \Pi(x) - x\Psi(0) = \sum_k \left(\frac{x}{k} - \log\frac{k+x}{k}\right), \tag{12}$$

oder für $x = 1$

$$-\Psi(0) = \sum_k \left(\frac{1}{k} - \log\frac{k+1}{k}\right). \tag{13}$$

Eliminiert man $\Psi(0)$ aus (12) und (13), so wird

$$\log \Pi(x) = \sum_k \left(x \log\frac{k+1}{k} - \log\frac{k+x}{k}\right),$$

oder, wenn man von den Logarithmen zu den Zahlen übergeht,

$$\Pi(x) = \left(\frac{1^{1-x}2^x}{x+1}\right)\left(\frac{2^{1-x}3^x}{x+2}\right)\left(\frac{3^{1-x}4^x}{x+3}\right)\cdots. \tag{14}$$

Diese Darstellung von $\Pi(x)$ durch ein unendliches und unbedingt konvergentes Produkt hat den Vorzug, daß sie für alle reellen und komplexen x gültig ist; sie lehrt, daß $\Pi(x)$ für endliche x niemals verschwindet, und für negative ganzzahlige x unendlich wird. Um die Art des Unendlichwerdens zu erkennen, bilde man

$$1 : (\Pi(x)\Pi(-x)) = \left(1 - \frac{x^2}{1}\right)\left(1 - \frac{x^2}{4}\right)\left(1 - \frac{x^2}{9}\right)\cdots$$

und beachte die bekannte Darstellung des Sinus, nämlich

$$\sin \pi x = \pi x\left(1 - \frac{x^2}{1}\right)\left(1 - \frac{x^2}{4}\right)\left(1 - \frac{x^2}{9}\right)\cdots.$$

Man erhält dann

$$\Pi(x)\Pi(-x) = \pi x : \sin \pi x \tag{15}$$

oder

$$\Pi(x-1)\Pi(-x) = \pi : \sin \pi x \tag{16}$$

und durch logarithmische Differentation von (15)

$$\Psi(x) - \Psi(-x) = \frac{1}{x} - \pi \cot \pi x. \tag{17}$$

Aus der Gleichung (15) ist zu entnehmen, daß, wenn x gegen die negative ganze Zahl $-n$ geht, das Produkt $\Pi(x)\sin \pi x$ den Wert $-\pi : \Pi(n-1)$ annimmt, wodurch die Art des Unendlichwerdens von $\Pi(x)$ vollständig bestimmt ist.

Wegen (3) ist es imgrunde genommen ausreichend, daß die Werte von $\Pi(x)$ in dem Intervall von $x = 0$ bis $x = 1$ bekannt sind. Dementsprechend hat *Gauß* (Gesammelte Werke, Band III, Seite 161) eine Tafel für den gemeinen Logarithmus von $\Pi(x)$ und für die logarithmische Ableitung $\Psi(x)$ gegeben, die mit dem Intervall 0.01 fortschreitet. Für größere Werte von x wird indessen die Anwendug jener Tafel in Verbindung mit (3) unbequem, und man hat dann andere Darstellungen zu benutzen.

§ 28. Setzt man in (16) und in (15) $x = \frac{1}{2}$, so wird

$$\Pi(-\tfrac{1}{2}) = \sqrt{\pi}, \qquad \Pi(\tfrac{1}{2}) = \tfrac{1}{2}\sqrt{\pi}. \tag{18}$$

Setzt man ferner in der Einführungsgleichung (1) für x den Wert $-\frac{1}{2}$ und führt statt t die Veränderliche u^2 ein, so wird

$$\Pi(-\tfrac{1}{2}) = \sqrt{\pi} = 2\int_0^\infty du \exp(-u^2). \tag{19}$$

Setzt man fest, daß die Nummer k von 1 bis ∞ laufen soll und trennt man in der aus (11) folgenden Gleichung

$$\Psi'(2x) = \sum_k (2x + k)^{-2}$$

die geraden Glieder von den ungeraden, so erhält man

$$\Psi'(2x) = \sum_k (2x + 2k - 1)^{-2} + \sum_k (2x + 2k)^{-2},$$

$$4\Psi'(2x) = \sum_k (x - \tfrac{1}{2} + k)^{-2} + \sum_k (x + k)^{-2}$$

$$= \Psi'(x - \tfrac{1}{2}) + \Psi'(x).$$

Integriert man diese Gleichung zweimal unbestimmt nach x und bezeichnet mit a und b die Integrationskonstanten, so wird

$$\log \Pi(2x) = \log \Pi(x - \tfrac{1}{2}) + \log \Pi(x) + ax + b.$$

Setzt man zunächst $x = 0$, so entsteht

$$b = -\log \Pi(-\tfrac{1}{2}),$$

setzt man ferner $x = 1$, so erhält man wegen (18) $a = \log 4$, woraus wenn man zu den Zahlen übergeht, die Beziehung

$$\Pi(2x)\,\Pi(-\tfrac{1}{2}) = 4^x \Pi(x - \tfrac{1}{2})\,\Pi(x) \tag{20}$$

fließt.

Ersetzt man in (9) u durch $\exp(-v)$, so wird

$$\Psi(x) - \Psi(0) = \int_0^\infty dv\, W, \qquad W = [1 - \exp(-xv)] : [\exp v - 1].$$

Differentiiert man darauf nach x, so entsteht

$$\Psi'(x) = \int_0^\infty dv\, V \exp(-xv), \qquad V = v : (\exp v - 1).$$

Setzt man endlich

$$V = 1 - \tfrac{1}{2}v + U, \qquad U = \frac{v}{2} \cdot \frac{\exp v + 1}{\exp v - 1} - 1, \tag{21}$$

so wird

$$\Psi'(x) = \int dv \exp(-xv) - \tfrac{1}{2}\int dv\, v \exp(-xv) + \int dv\, U \exp(-xv).$$

Die beiden ersten Integrale rechterhand lassen sich nach (5) unmittelbar auswerten, so daß

$$\Psi'(x) = x^{-1} - \tfrac{1}{2}x^{-2} + \int dv\, U \exp(-xv)$$

wird. Integriert man nunmehr zweimal unbestimmt nach x und bezeichnet mit A und B die Integrationskonstanten, so erhält man

$$\log \Pi(x) = x \log x - x + \tfrac{1}{2}\log x + Ax + B + J, \tag{22.a}$$

$$J = \int_0^\infty dv\, v^{-2} U \exp(-xv). \tag{22.b}$$

Hierin ist einerseits der Wert von A und B zu ermitteln, andererseits das Verhalten von J festzustellen.

§ 29. Der in (21) angesetzte Ausdruck U bleibt ungeändert, wenn man v durch $-v$ ersetzt, d. h. U ist eine gerade Funktion von v. Entwickelt man ferner U nach steigenden Potenzen von v, so fällt das konstante Anfangsglied heraus und man kann demgemäß die Reihe

$$U = B_1 \frac{v^2}{2!} - B_2 \frac{v^4}{4!} + B_3 \frac{v^6}{6!} - \cdots \tag{23}$$

ansetzen, wo die Faktoren B die Reihe der sogenannten *Bernoulli*schen Zahlen bilden, deren niedrigste Glieder sich unschwer durch direkte Entwicklung ergeben; man erhält so z. B.

$$B_1 = 1:6, \quad B_2 = 1:30, \quad B_3 = 1:42, \quad B_4 = 1:30, \quad \text{usw.}$$

Eine andere Darstellung von U entsteht, wenn man die Gleichung

$$\sin \pi x = \pi x \left(1 - \frac{x^2}{1}\right)\left(1 - \frac{x^2}{4}\right)\left(1 - \frac{x^2}{9}\right) \cdots$$

logarithmisch differentiiert und die auftretende Kotangente durch die Exponentialfunktion ausdrückt. Es wird dann

$$\pi i \frac{\exp(2i\pi x) + 1}{\exp(2i\pi x) - 1} = \frac{1}{x} + 2x \sum_k (x^2 - k^2)^{-1}, \quad (k = 1, 2, 3, \ldots).$$

Ersetzt man hierin $2i\pi x$ durch v, so wird mit Rücksicht auf (21)

$$U = 2v^2 \sum (v^2 + 4\pi^2 k^2)^{-1}.$$

Daraus fließt, wenn man wieder nach v entwickelt und die Potenzsummen

$$S_q = 1^{-q} + 2^{-q} + 3^{-q} + \cdots$$

einführt,

$$U = 2S_2 \left(\frac{v}{2\pi}\right)^2 - 2S_4 \left(\frac{v}{2\pi}\right)^4 + 2S_6 \left(\frac{v}{2\pi}\right)^6 - \cdots,$$

und durch Vergleichung mit (23)

$$(2\pi)^{2q} B_q = 2(2q)! S_{2q}. \tag{24}$$

Die für U angesetzte Partialbruchreihe bringt J in die Gestalt

$$J = 2\sum_k \int dv \exp(-xv)(v^2 + 4\pi^2 k^2)^{-1},$$

woraus mit der Substitution $v = 2\pi k w$ und der Abkürzung

$$\pi Z = \sum_k \frac{1}{k} \exp(-2\pi k x w) = -\log(1 - \exp(-2\pi x w))$$

für J die Darstellung

$$J = \int dw Z (1 + w^2)^{-1}$$

folgt. Setzt man jetzt für $(1 + w^2)^{-1}$ die Reihenentwicklung

$$(1 + w^2)^{-1} = 1 - w^2 + w^4 - \cdots \pm w^{2n-2} \mp w^{2n}(1 + w^2)^{-1}, \tag{25}$$

so liefert die Potenz w^{2m} zu J den Beitrag

$$\int dw\, w^{2m} Z = \sum_k \int \frac{dw}{\pi k} w^{2m} \exp(-2\pi k x w),$$

der mit Rücksicht auf (5) und (24) in

$$2(2m)!(2\pi)^{-2m-2} S_{2m+2} x^{-2m-1} = B_{m+1} : ((2m+1)(2m+2)x^{2m+1})$$

übergeht. Daraus folgt für J die Darstellung

$$J = \frac{B_1}{1 \cdot 2 \cdot x} - \frac{B_2}{3 \cdot 4 \cdot x^3} + \frac{B_3}{5 \cdot 6 \cdot x^5} - \cdots \pm \frac{B_n}{(2n-1)(2n)x^{2n-1}} + \text{Rest}. \tag{26}$$

Das Restglied besitzt den Ausdruck

$$\mp \int dw\, w^{2n} Z (1 + w^2)^{-1}.$$

Setzt man hierin für die negative Potenz von $1 + w^2$ einfach Eins, so erhält man statt des Restes einen zu großen numerischen Betrag, da Z niemals negativ ist. Andererseits besitzt der vereinfachte Ausdruck dieselbe Gestalt, wie die vorher untersuchten Glieder. Man hat also, wenn Θ einen positiven echten Bruch bedeutet,

$$\text{Rest} = \mp \Theta B_{n+1} : [(2n+1)(2n+2)x^{2n+1}]. \tag{27}$$

Es sind nunmehr noch die oben eingeführten Integrationskonstanten A, B zu ermitteln. Nach (22.a) ist

$$\log \Pi(x) = x \log x - x + \tfrac{1}{2}\log x + Ax + B + J(x), \tag{28}$$

wo bei J das Argument x der Deutlichkeit halber hinzugeschrieben ist. Setzt man hierin $x - 1$ für x und zieht die neue Gleichung von der alten ab, so entsteht nach der nötigen Reduktion unter Berücksichtigung von (3)

$$A = (x - \tfrac{1}{2})[\log(x-1) - \log x] + 1 + J(x-1) - J(x).$$

Auf der rechten Seite geht bei unbegrenzt wachsendem x der Bestandteil

$$(x - \tfrac{1}{2})[\log(x - 1) - \log x] + 1$$

gegen Null, und dasselbe gilt von den beiden J-Größen, wenn man die vorhin für $J(x)$ erhaltene Reihenentwicklung beachtet. Daraus ist zu entnehmen, daß $A = 0$ ist. Setzt man ferner in der aus (20) folgenden Gleichung

$$\log \Pi(2x) + \tfrac{1}{2}\log \pi = x \log 4 + \log \Pi(x - \tfrac{1}{2}) + \log \Pi(x)$$

für $\log \Pi(2x)$, $\log \Pi(x - \frac{1}{2})$ und $\log \Pi(x)$ die aus (28) folgenden Ausdrücke ein, so ergibt sich nach der nötigen Reduktion

$$B - \tfrac{1}{2}\log(2\pi) = x[\log x - \log(x - \tfrac{1}{2})] - \tfrac{1}{2} + J(2x) - J(x - \tfrac{1}{2}) - J(x).$$

Läßt man hierin x wieder über alle Grenzen wachsen, so ergibt sich ähnlich wie vorher, daß die rechte Seite null ist, woraus für B der Wert $\frac{1}{2}\log(2\pi)$ folgt. Daraus fließt schließlich für $\log \Pi(x)$ die Darstellung

$$\log \Pi(x) = x \log x - x + \tfrac{1}{2}\log(2\pi x) + J(x), \qquad (29)$$

in der für $J(x)$ der aus (26) und (27) folgende Ausdruck gesetzt zu denken ist. Die vorstehende Gleichung enthält die sogenannte *Stirling*sche Formel.

Um das Verhalten der Reihe (26) bei festem x und wachsendem n zu übersehen, hat man zunächst zu beachten, daß sich die in (24) auftretenden S-Größen für größere Werte ihres Index nicht weit von Eins entfernen. Daraus folgt nach (24) die genäherte Gleichung

$$B_{q+1} : B_q = (2q + 1)(2q + 2) : (2\pi)^2.$$

Die B-Zahlen wachsen also schließlich sehr rasch an und zwar derart, daß die Reihe (26), ins Unendliche fortgesetzt, sicher divergiert, wie groß man auch x gewählt haben möge. Andererseits nehmen für ein nicht zu kleines x die Reihenglieder anfangs ab, und da der Rest, wenn man bei einem bestimmten Gliede stehen bleibt, kleiner ausfällt als das Glied, das zunächst folgen würde, so ist die Reihe für größere x sehr wohl zur Berechnung von $J(x)$ brauchbar. Man hat es also im vorliegenden Falle mit einer sogenannten semikonvergenten Entwicklung zu tun.

§ 30. Die Funktion Π wird auch als *Euler*sches Integral zweiter Art bezeichnet, im Gegensatze zu den Integralen erster Art, die die Gestalt

$$\int_0^1 dt\, t^m (1 - t)^n$$

besitzen. Die Integrale erster Art lassen sich durch die Π-Funktionen ausdrücken. Zu dem Ende setzen wir in (1) u^2 für t und erhalten

$$\Pi(x) = 2\int_0^\infty du\, u^{2x+1} \exp(-u^2),$$

und ebenso

$$\Pi(y) = 2\int_0^\infty dv\, v^{2y+1} \exp(-v^2),$$

woraus

$$\Pi(x)\Pi(y) = 4\iint du\, dv\, u^{2x+1} v^{2y+1} \exp(-u^2 - v^2)$$

folgt. Denkt man sich nun u und v als rechtwinklige Koordinaten eines Punktes in der Ebene und führt die Polarkoordinaten r, s durch die Gleichungen

$$u = r\cos s, \qquad v = r\sin s$$

ein, so wird

$$\Pi(x)\Pi(y) = 4\iint dr\, ds\, r^{2x+2y+3} \cos s^{2x+1} \sin s^{2y+1} \exp(-r^2),$$

wo die Integration nach r von 0 bis ∞ und nach s von 0 bis $\frac{1}{2}\pi$ geht. Die Integration nach r läßt sich sofort ausführen, wenn t für r^2 gesetzt wird, und man erhält so zunächst

$$\Pi(x)\Pi(y) = 2\Pi(x+y+1)\int ds \cos s^{2x+1} \sin s^{2y+1}.$$

Setzt man weiter $\cos s^2 = t$, so wird

$$\Pi(x)\Pi(y) = \Pi(x+y+1)\int_0^1 dt\, t^x (1-t)^y, \tag{30}$$

womit die gesuchte Darstellung für die Integrale erster Art geleistet ist.

Fünfte Vorlesung.

Die Funktionen $\Phi(x)$ und $\operatorname{sg}(x)$.

§ 31. Außer der Transzendente $\Pi(x)$ kommt für die zu behandelnden Aufgaben noch eine andere mit $\Phi(x)$ bezeichnete Funktion in Betracht, welche durch die Gleichung

$$\sqrt{\pi}\,\Phi(x) = 2\int_0^x dt \exp(-t^2) \tag{1}$$

definiert ist. Setzt man für die Exponentialgröße ihre Reihenentwicklung ein und integriert, so wird

$$\sqrt{\pi}\,\Phi(x) = 2\left(\frac{x}{0!\,1} - \frac{x^3}{1!\,3} + \frac{x^5}{2!\,5} - \cdots\right). \tag{2}$$

$\Phi(x)$ ist hiernach durch eine ungerade und beständig konvergente Reihe dargestellt. Wird x unendlich, so erhält man durch Vergleichung von (1) mit der Formel (19) in § 28

$$\Phi(\infty) = 1 \quad \text{und} \quad \Phi(-\infty) = -1.$$

Da ferner die Ableitung von $\Phi(x)$ durch den stets positiven Ausdruck

$$2\exp(-x^2) : \sqrt{\pi}$$

gegeben ist, so geht $\Phi(x)$ beständig wachsend von -1 nach $+1$, wenn x von $-\infty$ nach $+\infty$ wandert.

Da die direkte Berechnung eines einzelnen Wertes von $\Phi(x)$ immerhin etwas umständlich ausfällt, so hat man die Funktion tabuliert, und es ist jetzt Gebrauch, den Werken über Wahrscheinlichkeitsrechnung eine solche Tafel als notwendiges Zubehör anzuhängen. Einige Notizen darüber findet man in dem Anhange am Schlusse dieses Werkes zusammengestellt. Ferner ist zu bemerken, daß die im Anhange abgedruckte vierstellige Tafel zu $\Phi(x)$ wesentlich für die Aufgaben der Kollektivmaßlehre bestimmt ist: man kommt bei diesen Aufgaben nur ganz ausnahmsweise in die Lage, in dem Werte von Φ mehr als vier Stellen berücksichtigen zu müssen. Infolgedessen hätte es keinen Sinn gehabt, in einer für den gewöhnlichen Gebrauch bestimmten Tafel noch die fünfte oder gar sechste Dezimale hinzuzufügen, die bei der Benutzung nur als ein recht hemmender Ballast wirken würden.

§ 32. Die Annäherung von $\Phi(x)$ an die obere Grenze $+1$ erfolgt sehr rasch, wie sich schon aus der Betrachtung einer mit größerer Stellenzahl gerechneten Tafel erkennen läßt. Da diese rasche Annäherung eine wichtige Rolle spielt, so untersuchen wir noch besonders das Verhalten von Φ für große Werte von x. Zu dem Ende führen wir ein

$$F(n) = \exp(x^2)\int_x^\infty dt \exp(-t^2)t^{-2n},$$

woraus zunächst

$$\sqrt{\pi}\,\Phi(x) = \sqrt{\pi} - 2\exp(-x^2)F(0) \tag{3}$$

folgt. Ferner erhält man aus

$$d[\exp(-t^2)t^{-2n-1}] : dt = -2\exp(-t^2)t^{-2n} - (2n+1)\exp(-t^2)t^{-2n-2}$$

durch Integration zwischen den Grenzen x und ∞

$$x^{-2n-1} = 2F(n) + (2n+1)F(n+1),$$

woraus für $n = 0, 1, 2, \ldots$ die Gleichungen

$$\begin{aligned} F(0) &= \tfrac{1}{2}x^{-1} - \tfrac{1}{2}F(1),\\ F(1) &= \tfrac{1}{2}x^{-3} - \tfrac{3}{2}F(2),\\ F(2) &= \tfrac{1}{2}x^{-5} - \tfrac{5}{2}F(3),\\ &\ldots\ldots\ldots \end{aligned}$$

fließen. Hiermit wird

$$F(0) = \frac{1}{2x} - \frac{1}{2^2x^3} + \frac{1\cdot 3}{2^3x^5} - \cdots \pm \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^{n+1}x^{2n+1}} \mp \text{Rest},$$

$$\text{Rest} = \frac{1\cdot 3\cdot 5\cdots(2n+1)}{2^{n+1}}F(n+1).$$

Setzt man nun in

$$F(n+1) = \exp(x^2)\int_x^\infty dt \exp(-t^2) t^{-2n-2}$$

für die Exponentialgröße unter dem Integralzeichen den größten Wert, den sie bei der Integration annimmt, nämlich $\exp(-x^2)$, so erhält man für $F(n+1)$ einen zu großen Wert. Man hat also, wenn man in dieser Weise verfährt und mit Θ einen gewissen positiven echten Bruch bezeichnet,

$$F(n+1) = \Theta x^{-2n-1} : (2n+1),$$

womit die Reihe für $F(0)$ die Gestalt

$$F(0) = \frac{1}{2x}\left(1 - \frac{1}{2x^2} + \frac{1\cdot 3}{(2x^2)^2} - \cdots \pm \frac{1\cdot 3\cdot 5\cdots(2n-1)}{(2x^2)^n}(1-\Theta)\right) \quad (4)$$

annimmt. Bildet man den Quotienten zweier aufeinander folgender Glieder, so erkennt man, daß die Reihe für hinlängliche große x semikonvergent ist, und daß der vernachlässigte Rest numerisch kleiner ist, als das letzte mitgenommene Glied.

Beschränkt man sich für große x auf das erste Glied der Reihe, so wird

$$\sqrt{\pi}(1-\Phi(x)) = \exp(-x^2) : x,$$

woraus man unmittelbar die rasche Abnahme von $1-\Phi$ für große x erkennt. Noch deutlicher wird dies aus folgendem Täfelchen:

$x = 1.0$	$1 - \Phi(x) = 0.157$
2.0	0.004 68
3.0	0.000 022 1
3.5	0.000 000 743
4.0	0.000 000 015 4.

§ 33. Um noch eine andere, weiterhin gebrauchte Darstellung von Φ zu erlangen, setzen wir an

$$K = \int_{-\infty}^{\infty} du \exp(-u^2)\cos(2xu)$$

und entwickeln den Kosinus nach Potenzen von x. Der Koeffizient von x^{2n} wird

$$\frac{(-4)^n}{(2n)!}\int du \exp(-u^2) u^{2n},$$

oder, wenn man unter Verdoppelung des Ausdruckes nur von 0 bis ∞ integriert und t für u^2 einführt,

$$(-4)^n \Pi(n-\tfrac{1}{2}) : \Pi(2n),$$

wofür nach (20) in § 28

$$(-1)^n \Pi(-\tfrac{1}{2}) : \Pi(n)$$

geschrieben werden kann. Fügt man jetzt den Faktor x^{2n} hinzu und summiert nach n, so wird

$$K = \sqrt{\pi} \exp(-x^2),$$

oder, wenn man die Ableitung von $\Phi(x)$ mit $\Phi(x)_1$ bezeichnet,

$$2\int_{-\infty}^{\infty} du \exp(-u^2) \cos(2xu) = \pi \Phi(x)_1 . \tag{5}$$

Integriert man nun beide Seiten nach x zwischen den Grenzen 0 und x, so erhält man

$$\int_{-\infty}^{\infty} \frac{du}{u} \exp(-u^2) \sin(2xu) = \pi \Phi(x) . \tag{6}$$

Setzt man hierin mit der positiven Konstante a für x den Quotienten $y:a$ und für u das Produkt av, so entsteht

$$\int \frac{dv}{v} \exp(-a^2 v^2) \sin(2yv) = \pi \Phi\left(\frac{y}{a}\right) .$$

Läßt man jetzt a gegen Null gehen, so wird auf der rechten Seite das Argument der Φ-Funktion, je nach dem Vorzeichen von y, positiv oder negativ unendlich, und man erhält

$$\int_{-\infty}^{\infty} \frac{dv}{v} \sin(2yv) = \pm \pi \quad \text{für} \quad y \gtrless 0 .$$

Würde man in dem vorstehenden Integral y von vornherein gleich Null setzen, so erhielte man offenbar statt des Wertes $\pm \pi$ den Wert Null. Um die drei hiernach möglichen Fälle kurz zusammenzufassen, führen wir das Zeichen $\operatorname{sg}(y)$ (gesprochen signum von y) ein, das den Wert $+1$ oder 0 oder -1 besitzen soll, je nachdem y positiv oder null oder negativ ist. Man hat dann

$$\pi \operatorname{sg}(y) = \int_{-\infty}^{\infty} \frac{dv}{v} \sin(2yv) . \tag{7}$$

Zur Erleichterung des späteren Gebrauches schreiben wir noch (6) und (7) unter Heranziehung der imaginären Einheit i in der Gestalt

$$\pi \Phi(x) = \int_{-\infty}^{\infty} \frac{dv}{iv} \exp(2ixv - v^2), \tag{8}$$

$$\pi \operatorname{sg}(x) = \int_{-\infty}^{\infty} \frac{dv}{iv} \exp(2ixv) . \tag{9}$$

Diese Umformung ist erlaubt, denn die unter dem Integralzeichen hinzugefügten imaginären Bestandteile sind ungerade, fallen also von selber heraus, sobald man zunächst zwischen den Grenzen $\pm C$ integriert und dann C unendlich werden läßt.

Außer der Funktion Φ werden auch ihre Ableitungen gebraucht,

die wir der Reihe nach mit $\Phi(x)_1, \Phi(x)_2, \ldots$ bezeichnen. Differentiiert man (8) p-mal nach x, so entsteht

$$\pi \Phi(x)_p = \int_{-\infty}^{\infty} \frac{dv}{iv} (2iv)^p \exp(2ixv - v^2). \tag{10}$$

Neben dieser Integraldarstellung ist noch eine andere vorhanden, zu der uns die aus (1) folgende Gleichnng

$$\sqrt{\pi}\, \Phi(x)_1 = 2 \exp(-x^2)$$

führt. Es ist zunächst

$$\begin{aligned} \sqrt{\pi}\, \Phi(x+v)_1 &= 2 \exp(-(x+v)^2) \\ &= 2\exp(-x^2 - 2xv - v^2) \\ &= \sqrt{\pi}\, \Phi(x)_1 \exp(-2xv - v^2). \end{aligned}$$

Entwickelt man den letzten Exponentialausdruck nach v in die Reihe

$$\exp(-2xv - v^2) = \sum_q \mathfrak{R}(x)_q (2v)^q, \quad (q = 0, 1, 2, \ldots), \tag{11}$$

so erhält man

$$\Phi(x+v)_1 = \Phi(x)_1 \sum \mathfrak{R}(x)_q (2v)^q,$$

also, wenn man links und rechts p-mal nach v differentiiert und dann $v = 0$ setzt,

$$\Phi(x)_{p+1} = 2^p p!\, \mathfrak{R}(x)_p \Phi(x)_1. \tag{12}$$

Das Bildungsgesetz der $\mathfrak{R}$ ergibt sich ohne Schwierigkeit, wenn man in (11) die beiden Ausdrücke

$$\exp(-2xv) \quad \text{und} \quad \exp(-v^2)$$

für sich nach v entwickelt und die beiden Reihen ausmultipliziert. Man erhält dann für die geraden $\mathfrak{R}$

$$\left.\begin{aligned} 2^0 \mathfrak{R}(x)_0 &= \frac{1}{0!\,0!} = 1, \\ 2^2 \mathfrak{R}(x)_2 &= \frac{(2x)^2}{0!\,2!} - \frac{1}{1!\,0!}, \\ 2^4 \mathfrak{R}(x)_4 &= \frac{(2x)^4}{0!\,4!} - \frac{(2x)^2}{1!\,2!} + \frac{1}{2!\,0!}, \\ &\ldots\ldots\ldots, \end{aligned}\right\} \tag{13}$$

und für die ungeraden $\mathfrak{R}$

$$\left.\begin{aligned} 2^1 \mathfrak{R}(x)_1 &= -\frac{2x}{0!\,1!}, \\ 2^3 \mathfrak{R}(x)_3 &= -\frac{(2x)^3}{0!\,3!} + \frac{2x}{1!\,1!}, \\ 2^5 \mathfrak{R}(x)_5 &= -\frac{(2x)^5}{0!\,5!} + \frac{(2x)^3}{1!\,3!} - \frac{2x}{2!\,1!}, \\ &\ldots\ldots\ldots, \end{aligned}\right\} \tag{14}$$

wo das Gesetz des Fortschreitens unmittelbar ersichtlich ist.

§ 34. Die $\mathfrak{R}_n$ sind gerade oder ungerade Funktionen, deren Ordnung mit der Nummer des betreffenden $\mathfrak{R}_n$ übereinstimmt. Beachtet man nun den Verlauf von $\Phi(x)_1$, so erkennt man aus (12), daß die Φ_p für große Werte von x in ähnlicher Weise gegen Null gehen, wie Φ_1, wenn auch die Abnahme mit wachsendem Index immer langsamer erfolgt. Werden ferner die Φ_p als Ordinaten zu der Abszisse x über derselben Abszissenachse abgetragen, so entspricht jedem Maximum oder Minimum einer Φ_p-Kurve bei dem nächstfolgenden Φ ein Schnitt mit der Abszissenaxe. Benutzt man den Satz, daß bei einer stetigen Kurve zwischen zwei Schnitten mit der Abszissenachse stets wenigstens ein Maximum oder Minimum liegt, so gelangt man beim Durchmustern der einzelnen Kurven zu dem Ergebnis, daß Φ_p im Endlichen p Maxima und Minima besitzt und die Abszissenachse in $p-1$ verschiedenen Punkten schneidet. Daraus folgt dann weiter, daß $\mathfrak{R}_p$ gerade p reelle und ungleiche Wurzeln besitzt, und daß die von den Wurzeln eingenommene Strecke mit zunehmendem Index der $\mathfrak{R}$ wächst.

Bezeichnet man die Ableitung eines $\mathfrak{R}$ mit $\mathfrak{R}'$ und bildet von (11) die Ableitung nach x, so wird

$$\exp(-2xv - v^2) = -\sum_q \mathfrak{R}'(x)_q (2v)^{q-1},$$

woraus durch Vergleichung mit (11) sofort

$$-\mathfrak{R}'(x)_q = \mathfrak{R}(x)_{q-1} \tag{15}$$

folgt. Ferner liefert die Ableitung von (12)

$$\Phi_{p+2} = 2^p p!\,\mathfrak{R}_p'\Phi_1 + 2^p p!\,\mathfrak{R}_p\Phi_2.$$

Drückt man hierin die Φ und $\mathfrak{R}'$ nach (12) und (15) aus, so gelangt man zu der Rekursionsformel

$$0 = (2p+2)\mathfrak{R}_{p+1} + 2x\mathfrak{R}_p + \mathfrak{R}_{p-1}, \tag{16}$$

aus der dann noch

$$0 = \Phi_{p+2} + 2x\Phi_{p+1} + 2p\Phi_p \tag{17}$$

folgt. Man kann (17) benutzen, um die höheren Φ aus den niedrigeren zu berechnen, indessen nimmt dabei für größere x oder p die Genauigkeit der Rechnung rasch ab, weil dann die Fehler der niedrigeren Φ vergrößert in die höheren eingehen.

Schreibt man die Gleichung (9) in der Gestalt

$$\begin{aligned}\pi\,\mathrm{sg}(y-x) &= \int \frac{dv}{iv}\exp(2iyv - 2ixv)\\ &= \int \frac{dv}{iv}\exp(2iyv - v^2)\cdot\exp(-2ixv + v^2),\end{aligned}$$

so läßt sich für den zweiten Exponentialfaktor nach (11) die Reihenentwickelung

$$\exp(-2ixv+v^2)=\sum\mathfrak{R}(x)_q(2iv)^q \tag{18}$$

ansetzen, womit

$$\pi\,\mathrm{sg}\,(y-x)=\sum\mathfrak{R}(x)_q\int\frac{dv}{iv}(2iv)^q\exp(2iyv-v^2) \tag{19}$$

wird. Hieraus fließt aber unter Berücksichtigung von (10) sofort die Reihenentwickelung

$$\pi\,\mathrm{sg}\,(y-x)=\sum_q\mathfrak{R}(x)_q\,\Phi(y)_q,\qquad (q=0,1,2,\ldots), \tag{20}$$

wo unter Φ_0 immer die Funktion Φ selber zu verstehen ist.

§ 35. Die vorstehend gefundene Reihe für sg$(y-x)$ wird uns weiterhin die Grundlage für die allgemeine Behandlung der Kollektivmaßlehre, sowie für gewisse mit den sogenannten großen Zahlen zusammenhängende Aufgaben der W.-R. liefern. Zu dem Ende ist noch die Konvergenzfrage zu erörtern, denn wenn auch die Entwickelung (18) für endliche v sicher konvergiert, so ist das doch nicht ohne weiteres ausreichend, weil bei der Integration in (19) auch unendlich große Werte von v in Betracht kommen. Da es mir nicht gelungen ist, die Frage nach der Konvergenz oder Divergenz der Reihe (20) in bündiger Weise zu entscheiden, so will ich einen Weg einschlagen, der die Benutzung von (20) zu interpolatorischen Darstellungen selbst dann rechtfertigt, wenn die Reihe im allgemeinen divergieren sollte.

Bildet man mit den konstanten Koeffizienten $b_0, b_1, \ldots$ den Ausdruck

$$Q=b_0\mathfrak{R}(x)_0+b_1\mathfrak{R}(x)_1+\cdots+b_n\mathfrak{R}(x)_n$$

und ordnet nach Potenzen von x, so entsteht das Polynom n-ten Grades

$$P=a_0+a_1x+\cdots+a_nx^n,$$

worin die Koeffizienten $a_0, a_1, \ldots$ homogene lineare Verbindungen der b sind. Umgekehrt läßt sich ein beliebiges Polynom P auf die Gestalt Q bringen, denn die aus (11) folgende Gleichung

$$\exp(-2xv)=\exp(v^2)\sum_q\mathfrak{R}(x)_q(2v)^q$$

lehrt, wenn links und rechts nach v entwickelt wird, daß die Potenz x^n linear homogen durch die Polynome $\mathfrak{R}_n, \mathfrak{R}_{n-2}, \mathfrak{R}_{n-4}, \ldots$ ausgedrückt werden kann. Bei dem Übergange von P auf Q erscheinen dann die Koeffizienten b als lineare homogene Verbindungen der Koeffizienten a.

Soll der Ausdruck Q für beliebige Werte von x verschwinden, so müssen die b sämtlich Null sein. Denn das Verschwinden von Q zieht die Gleichung $P=0$ nach sich; P kann jedoch für beliebige x

nur verschwinden, wenn die Koeffizienten a sämtlich null sind, woraus sofort das Verschwinden aller b folgt.

Nach dieser Vorbereitung bilden wir mit dem positiven echten Bruch t die Größen

$$w^2 = 1 - t^2, \qquad z = (y - xt) : w, \qquad Z = \Phi(z),$$

wo die Wurzelgröße w beständig positiv genommen werden soll. Da $\Phi(z)$ eine transzendente ganze Funktion von z ist, so kann, wenn man z als eine Funktion der drei Veränderlichen y, xt, t^2 ansieht, $\Phi(z)$ in eine Reihe von der Form

$$Z = \Phi(z) = \sum_m \sum_n \sum_p (m, n, p) y^m (xt)^n (t^2)^p \tag{21}$$

entwickelt werden, wo die Koeffizienten (m, n, p) reine Zahlengrößen sind, und die Summationen nach m, n, p von Null bis Unendlich laufen. Die Reihe ist sicher konvergent, so lange y und xt endlich sind, und so lange ferner der absolute Betrag von t unterhalb 1 bleibt. Ordnet man (21) nach t und setzt

$$Z = Z(0) + Z(1)t + Z(2)t^2 + \cdots, \tag{22}$$

so sind die $Z(n)$ Potenzreihen von x und y. Da in den einzelnen Gliedern von (21) der Exponent von x niemals höher werden kann, als der von t, so ist $Z(n)$ in x höchstens vom n-ten Grade, und man darf nach der vorhin gemachten Bemerkung ansetzen

$$Z(n) = Z(n, 0)\Re(x)_n + Z(n, 1)\Re(x)_{n-1} + Z(n, 2)\Re(x)_{n-2} + \cdots$$

oder kürzer

$$Z(n) = \sum_p Z(n, p)\Re(x)_{n-p}. \tag{23}$$

Hierin darf die Summation nach p von Null an bis ins Unendliche fortgesetzt werden, wenn man vorschreibt, daß die Zeichen $\Re(x)_q$ und $Z(n, p)$ die Null bedeuten sollen, wenn q negativ und p größer als n wird.

Geht man mit (23) in (22) ein, so wird

$$Z = \sum_n \sum_p Z(n, p) t^n \Re(x)_{n-p}, \qquad (n, p = 0, 1, 2, \ldots), \tag{24}$$

wo die Koeffizienten $Z(n, p)$ nur noch von y abhängen.

Bildet man von $Z = \Phi(z)$ die partiellen Ableitungen nach x und y, so ist

$$w \frac{\partial Z}{\partial y} = \Phi(z)_1, \qquad w \frac{\partial Z}{\partial x} = - t\Phi(z)_1,$$

woraus die Beziehung

$$t \frac{\partial Z}{\partial y} + \frac{\partial Z}{\partial x} = 0$$

folgt, Setzt man hierin die Reihe (24) ein und bezeichnet die q-te Ableitung von $Z(n, p)$ nach y mit $Z(n, p)_q$, so wird unter Berücksichtigung von (15)

$$0 = \sum_n \sum_p Z(n, p)_1 t^{n+1} \Re(x)_{n-p} - \sum_n \sum_p Z(n, p) t^n \Re(x)_{n-p-1}.$$

Spaltet man nach den Potenzen von t, so entstehen die Bedingungen

$$0 = \sum_p [Z(n, p)_1 - Z(n+1, p)] \Re(x)_{n-p},$$

in denen die Summation in Wirklichkeit nur von $p = 0$ bis $p = n$ auszudehnen ist. Diese Bedingungen können nach dem oben ausgesprochenen Satze für beliebige x nicht anders bestehen, als wenn die in [ ] eingeschlossenen Größen sämtlich verschwinden, d. h. es ist

$$Z(n+1, p) = Z(n, p)_1.$$

Setzt man hiernach die Gleichungenreihe

$$Z(n, p) = Z(n-1, p)_1, \quad Z(n-1, p) = Z(n-2, p)_1, \cdots Z(1, p) = Z(0, p)_1$$

an, so entsteht durch Elimination der Zwischengrößen die Beziehung

$$Z(n, p) = Z(0, p)_n. \tag{25}$$

Da nun die Größen $Z(0, 1)$, $Z(0, 2)$, ... identisch verschwinden, so gilt das Gleiche von $Z(n, 1)$, $Z(n, 2)$, ..., d. h. es sind nur die Größen $Z(n, 0)$ zu berücksichtigen. Da ferner für $t = 0$ der Ausdruck z in y übergeht, und dann unter Berücksichtigung von (22) und (23)

$$Z = \Phi(y) = Z(0) = Z(0, 0)$$

wird, so ergibt sich

$$Z(n, 0) = Z(0, 0)_n = \Phi(y)_n,$$

womit sich (24) in

$$Z = \Phi(z) = \sum_n t^n \Re(x)_n \Phi(y)_n \tag{26}$$

verwandelt. Schreibt man z in der Gestalt

$$z = \frac{y - xt}{w} = \frac{y - x}{w} + \frac{xw}{1 + t}$$

und läßt bei festgehaltenen x und y die Größe w gegen Null gehen, so geht z, je nachdem $y - x$ positiv oder negativ ist, gegen $+\infty$ oder $-\infty$, also $\Phi(z)$ gegen $+1$ oder -1, d. h. mit anderen Worten gegen sg $(y - x)$. Die Reihe (26) führt also, wenn $w = 0$ oder $t = 1$ gesetzt wird, auf die Gleichung (20). Da es jedoch einstweilen fraglich ist, ob die Reihe (26), die ja für $t < 1$ sicher konvergiert, diese Eigenschaft auch noch für $t = 1$ beibehält, so soll die gefundene Entwickelung in etwas anderer Weise nutzbar gemacht werden.

§ 36. In dem Ausdrucke $Z = \Phi(z)$ hängt z mit den Größen x, y, t, w durch die Gleichungen

$$wz = y - xt, \quad t^2 + w^2 = 1$$

zusammen. Ferner konvergiert Z gegen $\operatorname{sg}(y-x)$, wenn w gegen Null geht. Demnach läßt sich $\Phi(z)$ für kleine aber von Null verschiedene Werte der Größe w als eine genäherte Darstellung des Ausdruckes $\operatorname{sg}(y-x)$ ansehen, wobei der Fehler F dieser Näherung die Gestalt

$$F = \Phi(z) - \operatorname{sg}(y-x) \tag{27}$$

besitzt. Dies vorausgeschickt kann man fragen, ob sich der Fehler bei passender Wahl von w durchweg unter einer vorgeschriebenen Grenze halten läßt. Es wird sich zeigen, daß dies in der Tat möglich ist, sobald man dem Werte von $y-x$ eine gewisse Einschränkung auferlegt.

Statt x und y sollen zunächst die Größen X und Y durch die Gleichungen

$$2x = X - Y, \qquad 2y = X + Y$$

oder

$$y + x = X, \qquad y - x = Y$$

eingeführt werden, so daß

$$z = \frac{y - xt}{w} = \frac{(1+t)Y}{2w} + \frac{(1-t)X}{2w} \tag{28}$$

oder

$$z = \frac{(1+t)Y}{2w} + \frac{wX}{2(1+t)} \tag{29}$$

wird. Da ferner bei den späteren Anwendungen nur endliche Werte von x und y in Betracht kommen, so schreiben wir vor, daß X und Y ihrem numerischen Betrage nach stets unter einer bestimmten endlichen positiven Zahl C bleiben. Weiter sollen A und B zwei fest vorgeschriebene positive und beliebig kleine angebbare Zahlen bedeuten. Außerdem werde zu A aus der Gleichung

$$1 - A = \Phi(a) \tag{30}$$

die Zahl a ermittelt, wobei die früher gefundenen Gleichungen (3) und (4) benutzt werden können; die Zahl a besitzt dann einen bestimmten angebbaren wenn auch großen Wert. Endlich schreiben wir noch vor, daß Y vorläufig außerhalb des Intervalls mit den Grenzen $\pm B$ bleiben solle.

Beschränkt man Y zunächst auf positive Werte und läßt in (29) bei festgehaltenem w die Größen X, Y innerhalb der zulässigen Grenzen, d. h. X von $-C$ bis C und Y von B bis C variieren, so tritt der algebraisch kleinste Wert von z für die kleinsten Werte von X und Y, also für $X = -C$ und $Y = B$ ein und nimmt die Gestalt

$$z' = \frac{(1+t)B}{2w} - \frac{wC}{2(1+t)} \tag{31}$$

an. Läßt man hierin w von Null an wachsen, so sinkt z' von dem Anfangswerte $+\infty$ herab und wird, wenn w bis $+1$ geht, schließ-

lich sicher negativ, da ja für $w = 1$ und $t = 0$ der Ausdruck z' die Gestalt $(B - C) : 2$ annimmt. Man kann also einen Wert $w = G$ angeben, für den $z' = a$ wird. Läßt man demgemäß w nunmehr nicht über G hinausgehen, so sinkt der in (29) angesetzte Ausdruck z für die in Betracht kommenden Werte von X und Y sicher nicht unter a, entsprechend bleibt $\Phi(z)$ wegen (30) zwischen den Grenzen 1 und $1 - A$. Da nun im vorliegenden Falle sg $(y - x)$ oder sg (Y) den Wert $+1$ besitzt, so erscheint nach (27) der Fehler F in der Gestalt

$$F = -[1 - \Phi(z)]$$

und bleibt demzufolge zwischen den Grenzen 0 und $-A$.

Wiederholt man die vorstehende Betrachtung mutatis mutandis für den Fall eines negativen Y, so gelangt man zu dem Ergebnis, daß F zwischen den Grenzen 0 und A bleibt, wenn w wiederum nicht über G hinausgeht. Man kann also durch Einführung eines hinreichend kleinen Wertes von w bewirken, daß der Fehler F zwischen den beliebig eng zu wählenden Grenzen $\pm A$ liegt, so lange man für Y ein gewisses „kritisches" Intervall ausschließt, dessen Grenzen $\pm B$ ebenfalls beliebig eng gesteckt werden können.

Läßt man Y in das kritische Intervall rücken, so versagt die vorstehende Beweisführung, wie sich z. B. für den Fall $y = x$ oder $Y = 0$ ohne weiteres übersehen läßt. Der Fehler F kann dann die Grenzen $\pm A$ überschreiten, bleibt jedoch nach (27) stets zwischen den Grenzen ± 2, da weder sg $(y - x)$ noch $\Phi(z)$ dem numerischen Betrage nach über die Einheit hinausgehen.

Das vorstehende Ergebnis soll nun noch etwas abgeändert werden. Mit der endlichen Größe c und der angebbar von Null verschiedenen positiven Größe h bilden wir die Ausdrücke

$$u = h(x - c), \qquad v = h(y - c), \qquad z = (v - ut) : w, \tag{32}$$

wo t und w wieder die frühere Bedeutung haben sollen. Dann kann man, da sg $(v - u)$ wegen des positiven Vorzeichens von h mit sg $(y - x)$ zusammenfällt, die genäherte Darstellung

$$\operatorname{sg}(y - x) = \sum_q t^q \Re(u)_q \Phi(v)_q \tag{33}$$

ansetzen, deren Fehler die Gestalt

$$F = \sum_q t^q \Re(u)_q \Phi(v)_q - \operatorname{sg}(y - x) \tag{34}$$

annimmt. Da die Umformung (32) darauf hinauskommt, daß man für die Größen x, y den Anfangspunkt der Zählung und die zugrunde liegende Maßeinheit abändert, so läßt sich wiederum zeigen, daß bei Einführung eines genügend kleinen Wertes von w oder $1 - t$ der Fehler F zwischen den beliebig eng vorzuschreibenden Grenzen $\pm A$

bleibt, solange $y - x$ außerhalb eines beliebig eng vorzuschreibenden „kritischen" Intervalls mit den Grenzen $\pm B$ liegt, während innerhalb dieses Intervalls F zwischen den Grenzen ± 2 enthalten ist.

Der vorstehende Satz wird sich als ausreichend erweisen, um die später aufzustellende Fundamentalformel der Kollektivmaßlehre zu begründen.

Sechste Vorlesung.

Teilungsprobleme.

§ 37. Nach der Abschweifung in den beiden letzten Vorlesungen kehren wir wieder zu unserem eigentlichen Gegenstande zurück und wenden uns jetzt zu der Behandlung besonderer Aufgaben. Man kann hierbei zwei große Gruppen unterscheiden, je nachdem die Ursache der betrachteten Ereignisse bekannt ist oder nicht. Bei der ersten Gruppe, mit der wir uns zunächst beschäftigen wollen, kommt der direkte Weg zur Lösung darauf hinaus, daß man sich die Gesamtheit der möglichen Fälle zurechtlegt, dann die günstigen Fälle nebst ihren WW. heraussucht und daraus die W. des betrachteten Ereignisses ermittelt. Dieser direkte Weg ist jedoch häufig nicht der einzige; unter Umständen kann es vorteilhaft sein, zunächst gewisse Bedingungen aufzusuchen, denen die gesuchte W. zu genügen hat, und daraus die Lösung abzuleiten. Zur Erläuterung dieser Bemerkung sollen nachstehend zwei Teilungsprobleme behandelt werden. Als erstes Beispiel diene folgende Aufgabe.

A und B spielen gegen einander; das Spiel besteht aus einer Folge von Partien, von denen keine unentschieden bleibt; wer zuerst eine gewisse Anzahl von Partien gewonnen hat, gewinnt auch das Spiel; die W., eine Partie zu gewinnen, ist für A konstant gleich p, für B konstant gleich q, also $p + q = 1$; zum Gewinne des Spiels fehlen A und B noch x und y Partien; gesucht wird die W., daß A gewinnt.

Die gesuchte W. bezeichnen wir mit $W(x,y)$ und setzen $x + y - 1$ gleich n. Denkt man sich, ausgehend von dem augenblicklichen Stande des Spiels, eine beliebige Folge von Partien gespielt, so erkennt man, daß das Spiel unter allen Umständen mit spätestens n Partien beendet sein wird. Denn die stärkste Hinausschiebung des für A günstigen Ausgangs tritt ein, wenn B erst $y - 1$ Partien hintereinander gewinnt, worauf dann A die ihm fehlenden x Partien ebenfalls hintereinander zu gewinnen hat. Ferner sieht man, daß eine Partienfolge, die sich schon vor der n-ten Partie für A entschieden hat, nicht nachträglich bei der Fortsetzung bis zur n-ten Partie nun doch

noch von A verloren werden könnte. Demgemäß dürfen wir uns die Gesamtheit der möglichen Fälle in der Weise zurecht legen, daß wir alle Folgen von n Partien ansetzen und daraus die für A günstigen heraussuchen.

Die W. einer Partienfolge mit a Gewinnen für A und $n-a$ Gewinnen für B ist, wenn auf die Ordnung der Gewinne und Verluste keine Rücksicht genommen wird, nach den in § 18 gegebenen Formeln durch den Ausdruck

$$n!p^a q^{n-a} : a!(n-a)!$$

gegeben, der aus der Entwicklung von $(p+q)^n$ entspringt. Nun sind alle Folgen, in denen a nicht unterhalb x liegt, für A günstig, die anderen dagegen ungünstig. Man erhält also die gesuchte W. als Summe aller derjenigen Glieder in der Entwicklung von $(p+q)^n$, in denen der Exponent von p nicht unterhalb x liegt.

Die vorstehende, auf *Fermat* zurückgehende, Lösung ist offenbar äußerst einfach und läßt ohne weiteres die Ausdehnung auf den Fall von r Spielern $A_1, A_2, \ldots A_r$ zu. Ist bei dem Spieler A_h die W. für den Gewinn einer Partie gleich p_h, und fehlen ihm zum Gewinne des Spiels noch x_h Partien, so gelangt man durch Wiederholung der obigen Überlegungen zu dem Satze, daß die W. des für A_h günstigen Ausganges erhalten wird, wenn man in der Entwicklung von

$$(p_1 + p_2 + \cdots + p_r)^s, \quad (s = \sum_h x_h - r + 1)$$

alle Glieder summiert, in denen der Exponent von p_h nicht unterhalb x_h liegt.

§ 38. Zu dem Falle zweier Spieler zurückkehrend wollen wir jetzt die Lösung nach einem anderen Ansatze aufsuchen, dessen Grundgedanke auf *Pascal* zurückgeht. Wir zerlegen das betrachtete Ereignis E, nämlich Gewinn des Spieles durch A, in die beiden Arten E' und E'', je nachdem die erste der noch zu spielenden Partien von A gewonnen oder verloren wird. Nach dieser ersten Partie ist der Stand der noch fehlenden Partien $x-1$ zu y in dem Falle E' und x zu $y-1$ in dem Falle E''. Daraus erhält man nach der Produktregel, wenn $W(x, y)$ wieder die vorhin angegebene Bedeutung besitzt,

$$\mathfrak{W}(E') = p \cdot W(x-1, y), \quad \mathfrak{W}(E'') = q \cdot W(x, y-1), \tag{1}$$

also, weil nach der Summenregel $\mathfrak{W}(E) = \mathfrak{W}(E') + \mathfrak{W}(E'')$ ist,

$$W(x, y) = p\,W(x-1, y) + q\,W(x, y-1). \tag{2}$$

In der vorstehenden Gleichung müssen x und y zunächst mindestens gleich 2 sein, weil die sonst auf der rechten Seite auftretenden Zeichen

$W(0, y)$ und $W(x, 0)$ vorderhand keinen Sinn haben. Wir wollen nun den Wert dieser Zeichen so festsetzen, daß die Gleichungen (1) und (2) auch noch für $x = 1$ und $y = 1$ gültig bleiben. Für $x = 1$ ist die W., daß A das Spiel auf die Art E' gewinnt, gleich der W. für den Gewinn der ersten Partie, denn mit dem Gewinn dieser Partie hat A auch bereits das Spiel gewonnen. Also ist $\mathfrak{W}(E')$ gleich p, woraus nach (1) für $W(0, y)$ der Wert 1 folgt. Ist dagegen $y = 1$, so kann A auf die Art E'' das Spiel überhaupt nicht gewinnen, denn der Verlust der ersten Partie hat sogleich den Verlust des Spiels zur Folge. Daraus folgt nach der zweiten Gleichung in (1) für $\mathfrak{W}(E'')$ und $\mathfrak{W}(x, 0)$ der Wert Null.

Nach diesen Festsetzungen führen wir die beliebig klein zu wählenden Größen u und v ein, multiplizieren (2) mit $u^x v^y$ und summieren nach x und y von Eins bis Unendlich, woraus die Gleichung

$$\sum W(x, y) u^x v^y = \sum p W(x-1, y) u^x v^y + \sum q W(x, y-1) u^x v^y \qquad (3)$$

entspringt. Setzt man zur Abkürzung

$$S = \sum_x \sum_y W(x, y) u^x v^y, \qquad (x, y = 1, 2, 3, \ldots),$$

so erhält man in (3) für die linke Seite unmittelbar den Wert S. Das erste Glied rechts liefert

$$puS + \sum_y p W(0, y) u v^y = puS + \frac{puv}{1-v}.$$

Aus dem zweiten Gliede rechts entspringt

$$qvS + \sum_x q W(x, 0) u^x v = qvS,$$

womit (3) für S den Ausdruck

$$S = puv : [(1-v)(1-pu-qv)] \qquad (4)$$

liefert. Entwickelt man S wieder nach u und v, wozu man ja sehr verschiedene Wege einschlagen kann, so treten als Koeffizienten die gesuchten $W(x, y)$ auf. Damit ist offenbar zugleich ein Weg gewiesen, auf dem man die *Fermat*schen Summenausdrücke, die oben vorkamen, mannigfach umgestalten kann. Dieser Umstand kann als Beispiel für die Bemerkung dienen, daß die W.-R., ohne Rücksicht auf ihre praktischen Anwendungen, auch noch als eine selbständige Methode zur Entdeckung mathematischer Beziehungen Bedeutung hat.

Bei der Entwicklung von S will ich mich hier auf einen Weg beschränken, der für die $W(x, y)$ eine Integraldarstellung liefert. In der Gleichung

$$\int_0^\infty ds \exp(-as) = 1 : a, \qquad (a > 0) \qquad (5)$$

denken wir uns für a erst den einen, dann den anderen Faktor des Nenners von S eingesetzt und schreiben demgemäß

$$S = puv \int_0^\infty ds \exp[-(1-v)s] \int_0^\infty dt \exp[-(1-pu-qv)t]$$
$$= puv \iint ds\, dt \exp(-s-t) \exp(put) \exp(vs+qvt).$$

Der Ausdruck unter dem Integralzeichen läßt sich ohne weiteres nach u und v entwickeln, und man erhält für das gesuchte $W(x, y)$ die Gleichung

$$(x-1)!\,(y-1)!\,W = p^x \iint ds\, dt \exp(-s-t) t^{x-1} (s+qt)^{y-1},$$

also, wenn man erst nach s integriert und dabei die neue Veränderliche r durch $s = rt$ einführt,

$$= p^x \iint dt\, dr \exp(-t-tr) t^{x+y-1} (q+r)^{y-1},$$

woraus durch Integration nach t wegen Formel (5) in § 27

$$p^x (x+y-1)! \int dr\, (1+r)^{-x-y} (q+r)^{y-1}$$

entspringt. Hieraus folgt endlich mit der Substitution $1 + r = p : t$

$$(x-1)!\,(y-1)!\,W(x,y) = (x+y-1)! \int_0^p dt\, t^{x-1} (1-t)^{y-1}. \tag{6}$$

Entwickelt man den Ausdruck unter dem Integralzeichen nach Potenzen von t oder $1-t$, so erhält man eine nach Potenzen von p oder q fortschreitende Reihe, die indessen für größere Werte von x und y keinen besonderen Vorzug vor den Reihen besitzt, die man durch direkte Entwicklung von S erhalten kann. Dagegen läßt sich eine, gerade für große x, y bequeme Entwicklung nach den Φ-Funktionen vornehmen, mit der wir uns später bei dem Satze von *Bayes* zu beschäftigen haben werden.

Der benutzte Ansatz läßt sich verallgemeinern; es wird genügen dies an dem Falle zu zeigen, wo drei Spieler A, B, C vorhanden sind, denen zum Gewinne des Spiels x, y, z Partien fehlen, während die WW. für den Gewinn der einzelnen Partie die Werte p, q, r besitzen. Bedeutet $W(x, y, z)$ die W., daß A das Spiel gewinnt, so zerlege man das Ereignis E (Gewinn des Spiels durch A) in die drei Ereignisse E', E'', E''', je nachdem die erste der noch zu spielenden Partien von A oder B oder C gewonnen wird. Man erhält dann

$$\mathfrak{W}(E') = p \cdot W(x-1, y, z), \quad \mathfrak{W}(E'') = q \cdot W(x, y-1, z),$$
$$\mathfrak{W}(E''') = r \cdot W(x, y, z-1),$$

woraus die Differenzengleichung

$$W(x,y,z) = p\,W(x-1,y,z) + q\,W(x,y-1,z) + r\,W(x,y,z-1) \tag{7}$$

folgt. Ferner zeigt man wie oben, daß die Zeichen $W(0, y, z)$, $W(x, 0, z)$, $W(x, y, 0)$ die Werte $1, 0, 0$ besitzen. Multipliziert man nun (7) mit $u^x v^y w^z$ und summiert nach x, y, z von 1 bis ∞, so erhält man ähnlich wie oben

$$S = \sum_x \sum_y \sum_z W(x, y, z) u^x v^y w^z$$
$$= puvw : [(1 - v)(1 - w)(1 - pu - qv - rw)], \tag{8}$$

welcher Ausdruck, nach irgend einem passenden Verfahren entwickelt, zur Darstellung der $W(x, y, z)$ führt.

§ 39. Eine andere Aufgabe ist folgende. Die beiden Spieler A und B besitzen x und y Marken; bei jeder Partie erhält der Gewinnende eine Marke von dem Verlierenden; der Verlust aller Marken bedeutet zugleich den Verlust des Spiels; die WW., eine Partie zu gewinnen, sind für A und B konstant gleich p und q; gesucht wird die W., daß A das Spiel gewinnt.

Statt sogleich auf die gesuchte W. auszugehen, wollen wir zunächst die W. dafür ermitteln, daß A das Spiel mit gerade s Partien, aber nicht früher, gewinnt. Bezeichnet man diese W. mit $W(x, y, s)$ und unterscheidet bei dem Eintreten des zugehörigen Ereignisses die beiden Arten E' und E'', je nachdem die erste der zu spielenden Partien von A gewonnen oder verloren wird, so ist

$$\mathfrak{W}(E') = pW(x + 1, y - 1, s - 1), \quad \mathfrak{W}(E'') = qW(x - 1, y + 1, s - 1).$$

Die Summe der beiden W.-Größen gibt die W. des betrachteten Ereignisses, also ist

$$W(x, y, s) = pW(x + 1, y - 1, s - 1) + qW(x - 1, y + 1, s - 1). \tag{9}$$

Bei der Anwendung dieser Gleichung ist darauf zu achten, daß das Zeichen $W(x, y, s)$ vorläufig nur dann eine Bedeutung besitzt, wenn die beiden darin auftretenden Indizes x, y mindestens gleich Eins sind. Deshalb müssen in (9) wegen der auf der rechten Seite stehenden W.-Größen x und y mindestens gleich 2 sein. Die gleiche Bemerkung ist sinngemäß auf die späteren Formeln auszudehnen.

Das Zeichen $W(x, y, s)$ hängt außer von den Indizes x, y, s auch noch von den Größen p und q ab. Diese letzteren wollen wir zunächst herausschaffen. Man denke sich alle s-gliedrigen Kombinationen der beiden Buchstaben p, q angesetzt und lese jedes p als Gewinn, jedes q als Verlust einer Partie durch A. Dann entspricht jeder Kombination ein bestimmter möglicher Spielverlauf von s Partien, und umgekehrt jedem möglichen Spiele eine bestimmte Kombination. Liest man überdies jede Kombination als Produkt, so liefert dieses Produkt zugleich die W. für das Eintreten des zugehörigen Spielverlaufes.

Bei der Aufsuchung von $W(x, y, s)$ kommen nur solche Kombinationen in Betracht, die für B nach Beendigung der letzten Partie

mit dem Markenbesitz Null abschließen. Zu dem Ende muß A erstlich die y Marken gewinnen, die B anfangs besaß, dann aber auch jede Marke, die etwa im Verlaufe des Spiels an B verloren ging, wieder zurückzugewinnen. Auf k an B verlorene Partien müssen also $y + k$ gewonnene kommen, so daß

$$s = k + (y + k) = y + 2k$$

ist. Für jedes dieser Spiele wird die W. des Eintretens gleich dem Produkt $q^k p^{y+k}$ oder gleich $p^y (pq)^k$.

Die Anzahl der Kombinationen, die auf das vorstehende Produkt führen, ist gleich dem Binomialkoeffizienten

$$(y + 2k)! : [k! \, (y + k)!].$$

Da jedoch wegen der Bedeutung von $W(x, y, s)$ alle Kombinationen zu streichen sind, bei denen sich das Spiel schon vor der s-ten Partie entscheidet, so haben wir anzusetzen

$$W(x, y, s) = U(x, y, k) p^y (pq)^k, \quad (s = y + 2k), \qquad (10)$$

wo $U(x, y, k)$ eine gewisse ganze Zahl bedeutet, die von den Indexwerten x, y, k abhängt und im allgemeinen unterhalb jenes Binomialkoeffizienten liegt. Der Index k bedeutet hierbei die Zahl der von A verlorenen Partien und kann daher die Reihe der Zahlen $0, 1, 2, \ldots$ durchlaufen, sobald wir der Zahl s nach und nach alle zulässigen Werte erteilen. Ist $k = 0$ oder $s = y$, so gewinnt A hintereinander sämtliche Marken von B; die W. für ein solches Spiel ist gleich p^y, also

$$W(x, y, y) = p^y, \qquad U(x, y, 0) = 1. \qquad (11)$$

Ferner ist für $x = 1$, $y = 1$ das Spiel unter allen Umständen schon mit der ersten Partie entschieden, daher ist, außer für $k = 0$, die Größe $U(1, 1, k)$ beständig gleich Null.

§ 40. Setzt man jetzt in (9) die Ausdrücke (10) ein, so fallen die p, q heraus, und man erhält

$$U(x, y, k) = U(x + 1, y - 1, k) + U(x - 1, y + 1, k - 1). \qquad (12)$$

Summiert man die Gleichung (10) nach k von $k = 0$ an, so erhält man links die W. dafür, daß A überhaupt mit irgend einer Anzahl von Partien gewinnt. Da diese W. einen bestimmten endlichen Wert besitzt, so ist die unendliche Reihe, die man auf der rechten Seite erhält, sicher konvergent. Deshalb ist auch die mit einer beliebigen Veränderlichen t angesetzte Reihe

$$V(x, y) = \sum_k U(x, y, k) t^k, \quad (k = 0, 1, 2, \ldots) \qquad (13)$$

sicher konvergent, so lange t nicht über pq hinausgeht. Das Anfangsglied dieser Reihe ist nach (11) jedesmal gleich Eins, ferner reduziert sich $V(1, 1)$ nach der vorhin über die $U(1, 1, k)$ gemachten Bemerkung auf den konstanten Wert Eins.

Multipliziert man (12) mit t^k und summiert nach k von Eins an, so erhält man, nach der über das Anfangsglied der V gemachten Bemerkung,

$$tV(x-1, y+1) = V(x, y) - V(x+1, y-1). \tag{14}$$

Die vorstehend auftretenden V-Größen beziehen sich auf Spiele, bei denen der anfängliche Markenstand entweder $x-1$ zu $y+1$ oder x zu y oder endlich $x+1$ zu $y-1$ ist. Die Markensumme hat also in allen diesen Spielen denselben Wert $x+y$, wofür wir kurz z schreiben wollen. Schreibt man ferner bei konstantem z für den Augenblick statt $V(x, y)$ einfach $V(y)$, so nimmt die Gleichung (14) die Gestalt

$$tV(y+1) = V(y) - V(y-1) \tag{14.a}$$

an. In der Lehre von den Differenzengleichungen wird nun gezeigt, wie man Bedingungen der Form (14.a) zu behandeln hat. Man hat zunächst zwei voneinander unabhängige Partikularlösungen zu suchen, d. h. zwei Ausdrücke $F(y)$ und $G(y)$, die für $V(y)$ eingesetzt die vorgelegte Bedingung befriedigen und sich nicht bloß um einen konstanten Faktor unterscheiden. Sind diese gefunden, so hat man des weiteren die gesuchte Lösung in der Gestalt

$$V(y) = AF(y) + BG(y)$$

anzusetzen, wo die Koeffizienten A, B zwei von y unabhängige Größen bedeuten. Endlich sind diese Koeffizienten so zu bestimmen, daß der angesetzte Ausdruck auch den anderweitigen für $V(y)$ geltenden Bedingungen genügt.

Um die verlangten Partikularlösungen zu erhalten, suchen wir zu der quadratischen Gleichung

$$tX^2 - X + 1 = 0$$

die beiden Wurzeln u und v auf und setzen die hieraus entspringenden Beziehungen

$$2tu = 1 + \sqrt{(1-4t)}, \quad 2tv = 1 - \sqrt{(1-4t)},$$
$$tu + tv = 1, \quad tuv = 1$$

an. Dann sind, wie man sich leicht überzeugt, die Ausdrücke

$$F(y) = u^y, \quad G(y) = v^y$$

Lösungen von (14.a), so daß man, wenn man wieder zu der Bezeichnung $V(x, y)$ zurückkehrt,

$$V(x, y) = Au^y + Bv^y \tag{15}$$

setzen darf, wo nunmehr die A, B zu ermitteln sind. Zu dem Ende haben wir uns nach den weiteren Eigenschaften der Größen $W(x, y, s)$ umzusehen.

§ 41. Wir betrachten jetzt die Spiele, die unter Festhaltung der Markensumme z zu

$$W(z-1,\ 1,\ s) = W(z-1,\ 1,\ 1+2k) = U(z-1,\ 1,\ k)\,p(pq)^k$$

gehören, wobei jedoch die bereits früher erledigten Fälle $z = 2$ und $k = 0$ beiseite bleiben sollen. Die zur Darstellung des Spielverlaufes benutzten Kombinationen aus den Buchstaben p, q müssen dann sämtlich mit q anfangen, weil sonst das Spiel, wider die gemachte Voraussetzung, schon mit der ersten Partie zu Ende wäre. Faßt man nun diese Kombinationen in Gruppen zusammen, je nachdem sie mit qp oder qqp oder $qqqp$ usw. anfangen, so bedeuten diese Gruppen ebensoviele verschiedene Arten für das Zustandekommen von $W(z-1,\ 1,\ s)$ und wir dürfen darnach, indem wir ähnlich wie bei der Herstellung der Gleichung (9) vorgehen, ansetzen

$$\begin{aligned} W(z-1, 1, s) &= qp\,W(z-1, 1, s-2) + q^2p\,W(z-2, 2, s-3) + \cdots \\ &= \sum_n q^n p\,W(z-n,\ n,\ s-n-1) \end{aligned}$$

oder, wenn wir auf die U-Größen übergehen,

$$U(z-1,\ 1,\ k) = \sum_n U(z-n,\ n,\ k-n).$$

Der Summationsindex n bedeutet hierbei die Anzahl der Partien, die bei dem n-ten betrachteten Spielverlauf von A anfangs verloren werden. Daher darf n nicht über $z-2$ hinausgehen, weil andernfalls A dabei seinen ganzen Markenbesitz und damit auch das Spiel verlöre. Da ferner k die Anzahl der Marken bedeutet, die A bei den betrachteten Spielen überhaupt verliert (und wiedergewinnt), so darf n auch nicht über k hinausgehen, d. h. n darf nur bis zu der kleineren der beiden Zahlen $z-2$ und k gehen. Um diese beiden Fälle nicht besonders auseinander halten zu müssen, führen wir das Zeichen (a, b) ein, das den Wert Eins oder Null bedeuten soll, je nachdem b unterhalb a liegt oder nicht. Damit schreiben wir

$$U(z-1,\ 1,\ k) = \sum_n U(z-n,\ n,\ k-n)\cdot(z-1,\ n)\cdot(k+1,\ n) \qquad (16)$$

und dürfen nunmehr n von 1 bis ∞ laufen lassen, da die unzulässigen Summenglieder durch die Hinzufügung der Faktoren $(z-1,\ n)$ und $(k+1,\ n)$ unschädlich gemacht werden. Ferner sieht man, daß bei dieser Schreibweise nun auch der bisher ausgeschlossene Wert $z = 2$ zulässig wird.

Multipliziert man (16), das für $k = 1, 2, \ldots$ gilt, mit t^k und summiert nach k von 1 an, so wird wegen (13)

$$V(z-1,\ 1) - 1 = \sum_n (z-1,\ n)\,t^n \sum_k U(z-n,\ n,\ k-n)\cdot(k+1,\ n)\,t^{k-n}.$$

Summiert man auf der rechten Seite zuerst nach k und führt den neuen Summationsbuchstaben h durch die Bedingung $k = n + h$ ein, so darf man wegen des Verhaltens des Faktors $(n + h + 1, n)$ die Summation mit $h = 0$ beginnen und erhält

$$\begin{aligned} V(z-1, 1) - 1 &= \sum_n (z-1, n) t^n \sum_h U(z-n, n, h) t^h \\ &= \sum_n (z-1, n) t^n V(z-n, n), \end{aligned}$$

wofür wir jetzt

$$V(z-1, 1) - 1 = \sum_n t^n V(z-n, n), \quad (n = 1, 2, \ldots z-2)$$

schreiben wollen. Hierin setzen wir den aus (15) folgenden Ausdruck

$$V(z-n, n) = A(z) u^n + B(z) v^n$$

ein und erhalten nach gehöriger Reduktion

$$A(z)[t + (tu)^z] + B(z)[t + (tv)^z] = t. \tag{17}$$

§ 42. Um eine weitere Bedingung für die A, B zu finden, untersuchen wir noch die zu

$$W(1, z-1, s) = W(1, z-1, z-1+2k) = U(1, z-1, k) p^{z-1} (pq)^k$$

gehörigen Spiele, wobei wir wieder die Fälle $k = 0$ und $z = 2$ beiseite lassen. Die in Betracht zu ziehenden Kombinationen der p, q fangen dann sämtlich mit p an. Wir gruppieren nach den Spielanfängen pq, ppq, $pppq$, usw. und schreiben demgemäß

$$\begin{aligned} W(1, z-1, s) &= pq W(1, z-1, s-2) + p^2 q W(2, z-2, s-3) + \cdots \\ &= \sum_n p^n q W(n, z-n, s-n-1), \end{aligned}$$

$$U(1, z-1, k) = \sum_n U(n, z-n, k-1),$$

wobei n von 1 bis $z - 2$ zu laufen hat, da der Fall $n = z - 1$ auf ein Spiel führt, das schon vor der s-ten Partie für A entschieden ist. Multipliziert man die letzte Gleichung mit t^k und summiert nach k von 1 an, so wird

$$V(1, z-1) - 1 = \sum_n t V(n, z-n).$$

Die Einsetzung von

$$V(n, z-n) = A(z) u^{z-n} + B(z) v^{z-n}$$

liefert nach gehöriger Reduktion

$$A(z)[1 + tu^z] + B(z)[1 + tv^z] = 1. \tag{18}$$

Eliminiert man nun A und B aus (15), (17) und (18), so erhält man die fertige Lösung

$$V(x, y) = [(tu)^x - (tv)^x] : [(tu)^z - (tv)^z]. \tag{19}$$

Bezeichnet man mit $W(A)$ und $W(B)$ die WW., daß A oder B überhaupt gewinnt, so ist

$$W(A) = \sum_k W(x,\, y,\, y + 2k) = \sum_k U(x,\, y,\, k) p^y (pq)^k,$$

wo die Summation nach k von 0 bis ∞ geht. Hiernach ist

$$W(A) = p^y V(x,\, y),$$

vorausgesetzt, daß in V anstatt t das Produkt pq eingeführt wird. In diesem Falle wird aber

$$1 - 4t = 1 - 4pq = (p + q)^2 - 4pq = (p - q)^2.$$

Da in (19) die Größen u und v symmetrisch auftreten, so ist es gleichgültig, welches Vorzeichen man der Wurzel aus $1 - 4t$ gibt. Wir dürfen daher ansetzen

$$2tu = 1 + (p - q) = 2p, \qquad 2tv = 1 - (p - q) = 2q.$$

Damit wird

$$W(A) = p^y (p^x - q^x) : (p^z - q^z),$$

woraus durch Buchstabenvertauschung

$$W(B) = q^x (p^y - q^y) : (p^z - q^z)$$

folgt. Da hiernach $W(A) + W(B)$ gleich Eins wird, so ist die W. einer unbegrenzten Dauer des Spiels null, d. h. genauer ausgedrückt unendlich klein. Ist ferner $p = q = 0.5$, so wird

$$W(A) = x : z, \qquad W(B) = y : z,$$

d. h. die Gewinnaussichten sind in jedem Augenblicke den beiden Markenbeständen proportional.

Siebente Vorlesung.

Die Spieleinsätze.

§ 43. Die in der letzten Vorlesung behandelten Beispiele gehören zur Klasse der sogenannten Teilungsprobleme. Um diese Bezeichnung verständlich zu machen, erinnere ich zunächst an den in der Einleitung erwähnten Vorfall, der in den Händen von *Pascal* und *Fermat* zur Entstehung der W.-R. geführt hat. Es handelte sich dort um die Frage, wie die Einsätze eines Glückspiels zu verteilen seien, wenn das Spiel vor seiner Beendigung abgebrochen wird, wenn also, um es kurz auszudrücken, das Endspiel unterbleibt. Die Antwort, die wir an der früheren Stelle nur andeuten konnten, läßt sich jetzt an der Hand der bisherigen Entwicklungen folgendermaßen for-

mulieren. Man berechne die WW. des Gewinnens, die für die einzelnen Spieler bei dem betrachteten Endspiel vorhanden sind, und verteile die verfügbare Summe proportional zu jenen W-Größen. Um die Berechtigung dieser Verteilungsart einzusehen, denke man sich, daß für den Spieler A die W. des Gewinnens z. B. gleich 0.9 war, dann hatte er danach die Aussicht, bei wiederholter Ausführung des Endspiels unter zehn Fällen neunmal zu gewinnen; deshalb erscheint es als billig, daß ihm neun Zehntel der zu verteilenden Summe zugesprochen werden. Hierdurch wird die übliche Benennung der in der letzten Vorlesung behandelten beiden Beispiele verständlich: die genannten Aufgaben sind als Endspiele zu denken, bei denen behufs Verteilung der Einsätze die WW. des Gewinnens berechnet werden. In dieser Beziehung ist namentlich die zweite Aufgabe von Interesse, weil sich bei ihr die wirkliche Beendigung des Spiels unter Umständen recht lange hinausziehen kann.

§ 44. Soll ein Glücksspiel von Anfang an den Forderungen der Billigkeit genügen, so müssen die Einsätze und die Gewinnaussichten dem entsprechend gegeneinander abgewogen sein. Wenn z. B. nur um die Einsätze gespielt wird, und die Gewinnchancen der Spieler einander gleich sind, so leuchtet es ohne weiteres ein, daß auch Gleichheit der Einsätze stattzufinden hat. Wenn ferner die Chance von A zehnmal größer ist, als die von B, so ist es von vornherein plausibel, daß A diesen Unterschied durch einen zehnmal größeren Einsatz auszugleichen hat. Solche einfachen Überlegungen reichen jedoch nicht aus, wenn die Spielbedingungen verwickelter sind, und die Gewinne und Verluste, je nach dem Verlaufe des Spiels, Abstufungen erfahren. Um über solche Fälle Klarheit zu erlangen, schlagen wir folgenden Weg ein.

Zunächst setzen wir fest, daß Verluste als negative Gewinne gerechnet werden sollen, und verstehen unter Gewinn oder Verlust des Spielers A die positive oder negative Änderung, die das Vermögen von A infolge des Spiels erfährt. Ferner seien a, a', a'', ... die Vermögensänderungen, die für A innerhalb der Gesamtheit der möglichen Verläufe eines betrachteten Spieles in Frage kommen, wobei unter den a auch der Wert Null auftreten kann; entsprechend bezeichnen wir mit $W(a)$, $W(a')$, $W(a'')$, ... die WW., mit denen man das Eintreten der genannten a zu erwarten hat. Dies festgesetzt suchen wir die Vermögensänderung auf, die zu erwarten ist, wenn das Spiel mit ungeänderten Spielbedingungen n-mal ausgeführt wird, wobei wir uns n so groß denken, daß eine hinreichende Ausgleichung der dem einzelnen Spiele anhaftenden Zufälligkeiten stattfindet.

Bei einem hinreichend großen n ist die Häufigkeit, mit der die Vermögensänderung a erwartet und näherungsweise auch tatsächlich verwirklicht wird, durch das Produkt $nW(a)$ gegeben. Daraus ent-

springt als Beitrag zu der gesamten Vermögensänderung das Produkt $an W(a)$. Ähnliche Beiträge liefern die Größen a', a'', ..., so daß die Gesamtänderung die Gestalt

$$nh = an W(a) + a' n W(a') + \cdots$$

annimmt, wo zur Abkürzung

$$h = a W(a) + a' W(a') + \cdots \qquad (1)$$

gesetzt ist. Den Ausdruck h nennen wir die *totale mathematische Hoffnung* des Spielers A; entsprechend bezeichnen wir die Produkte $a W(a)$, ... als die, positiv oder negativ anzusetzenden, *partialen* Hoffnungen. Die Größe h gibt offenbar die Vermögensänderung an, die A — *im Durchschnitt aus einer hinreichend großen Anzahl von Spielen* — von dem einzelnen Spiel zu erwarten hat. Hierbei ist der Zusatz, daß es sich um einen Durchschnitt aus vielen Spielen handelt, sehr wesentlich, wie wir später noch des näheren erkennen werden.

Die vorstehende Berechnungsweise von h kann als der *Gewinnansatz* bezeichnet werden. Eine andere Berechnungsweise, die als der *Zahlungsansatz* zu bezeichnen ist, erhält man, wenn man statt der positiven oder negativen Gewinne die Zahlungen betrachtet, die von A bei dem Spiele, je nach seinem Verlaufe, geleistet oder empfangen werden. Es seien z, z', ... die Zahlungen, die bei den verschiedenen Spielverläufen in Frage kommen können, wobei wir die z positiv oder negativ rechnen, je nachdem A empfängt oder gibt, ferner seien $W(z)$, $W(z')$, ... die WW., die den einzelnen z zugehören; dann gelangt man, wenn wieder die aus n Spielen zu erwartende Vermögensänderung aufgesucht wird, zu dem Ansatze

$$h = z W(z) + z' W(z') + \cdots. \qquad (2)$$

Der Unterschied zwischen den beiden Ausdrücken für h läßt sich am besten an einem einfachen Beispiel deutlich machen. A und B spielen gegeneinander mit den Gewinnchancen $W(A)$ und $W(B)$, wobei

$$W(A) + W(B) = 1$$

sein soll, ferner seien x und y die bei Beginn des Spiels von A und B zu leistenden Einsätze, deren Summe dem Gewinner zufließt, dann sind für A die beiden Vermögensänderungen $+y$ und $-x$ möglich, und man erhält

$$h = y W(A) - x W(B).$$

Mit den z dagegen stellt sich die Rechnung folgendermaßen. Zunächst bedeutet für A der von ihm geleistete Einsatz eine negative Zahlung $z = -x$, deren $W(z)$ gleich Eins ist, da ja der Einsatz unter allen Umständen zu zahlen ist. Gewinnt nun A, so tritt die

positive Zahlung $z = x + y$ auf, deren $W(z)$ gleich $W(A)$ ist. Verliert A, so empfängt A die Zahlung $z = 0$, deren $W(z)$ gleich $W(B)$ ist. Also wird

$$\begin{aligned} h &= -x \cdot 1 + (x + y) \cdot W(A) + 0 \cdot W(B) \\ &= y W(A) - x W(B), \end{aligned}$$

was mit dem vorigen Ausdrucke, wie es sein soll, übereinstimmt.

Der erste Ansatz für h beruht darauf, daß man die Gesamtheit der möglichen Spielverläufe nach dem Betrage der eintretenden Vermögensänderung in gewisse, einander ausschließende, Gruppen zusammenfaßt, so daß die Summe der $W(a)$ den Wert Eins besitzt, während die Summe der $W(z)$, wie auch das gegebene Beispiel erkennen läßt, im allgemeinen von Eins verschieden ist.

Statt der Ausdrücke „positive oder negative mathematische Hoffnung" finden sich auch die Bezeichnungen „mathematische Hoffnung oder Furcht", ferner „Vorteil oder Nachteil", sowie „positive oder negative Erwartung".

§ 45. Die Bedingungen eines Spieles sind, im Durchschnitt aus hinreichend vielen Wiederholungen, für A günstig oder ungünstig, je nachdem das für A geltende h positiv oder negativ ist. Da nun der Nutzen oder Schaden von A für seinen Gegenspieler das Umgekehrte bedeutet, so muß, wenn keiner der Spieler bevorzugt sein soll, für jeden die Größe h gleich Null sein. Daraus folgt z. B. für den oben betrachteten einfachen Fall, daß sich die Einsätze umgekehrt wie die Gewinnchancen verhalten müssen, wenn das Spiel als nach Billigkeit geordnet gelten soll.

Die Bedingung $h = 0$ bildet den Ausgangspunkt für die Aufstellung der Spielbedingungen namentlich in solchen Fällen, wo ein einzelner „Spielhalter" gegen eine größere Anzahl von Einzelspielern oder Teilnehmern das Spiel „hält", wie das z. B. bei den Spielbanken und Lotterien zutrifft. Allerdings dient jene Bedingung nur als Ausgangspunkt, indem die daraus fließenden Beziehungen zwischen den Einsätzen, den Chancen und den Gewinnen bei der Ausführung nicht unerheblich zugunsten des Spielunternehmers „verbessert" werden. Denn dieser betreibt die Sache als ein Geschäft und will nicht nur die manchmal recht bedeutenden Geschäftsunkosten decken, sondern auch einen Unternehmergewinn erzielen, der so groß sein soll, als das ohne Lähmung des Anreizes zum Spielen nur immer angängig ist. Außerdem treffen die Spielbanken noch eine andere Vorkehrung, die in ihrer Gesamtwirkung ebenfalls auf den Vorteil des Unternehmers berechnet ist. Um dies deutlich zu machen betrachten wir ein einfaches Beispiel.

Man denke sich ein Spiel, bei dem Bankhalter und Einzelspieler mit gleichen Einsätzen und gleichen Gewinnchancen gegeneinander

spielen. Ferner gelte die Vorschrift, daß bezüglich der Höhe der Einsätze und der Dauer des Spieles lediglich das Belieben des Einzelspielers maßgebend sei. Dann würde diese Bedingung zum Nachteile des Bankhalters ausschlagen können, sobald der Einzelspieler über hinreichende Mittel verfügt oder sobald sich mehrere Einzelspieler zusammentun. Der Einzelspieler hätte zu dem Ende nur nötig, nach folgender Regel vorzugehen. Er setzt nach jeder gewonnenen Partie den Einsatz Eins, nach jeder verlorenen dagegen das Doppelte des eben verlorenen Einsatzes, setzt also bei einer Reihe von ungünstigen Partien nach der Skala 1, 2, 4, 8, ..., bis die erste günstige Partie, wie man leicht erkennt, ihm nicht nur die vorangegangenen Verluste wieder einbringt, sondern auch noch einen Gewinn Eins liefert. Offenbar wird die für den Bankhalter gefährliche Ausnutzung einer solchen Einsatzsteigerung verhindert, wenn von ihm, wie das Regel ist, für die Höhe der Einsätze eine gewisse untere und obere Grenze vorgeschrieben wird.

§ 46. Die Berechnung von h ist bei manchen der verbreiteteren Glücksspiele ganz interessant, indessen will ich darauf nicht eingehen, sondern eine Aufgabe behandeln, die in der Geschichte der W.-R. eine berühmte oder, wenn man lieber will, eine berüchtigte Rolle gespielt hat. Es ist dies das sogenannte Petersburger Problem.

A spielt gegen B das Spiel „Bild oder Schrift". Das Spiel ist auf n Würfe verabredet; A zahlt an B den Einsatz x, der in den Händen von B verbleibt; dafür zahlt B an A die Beträge 2, 4, 8, 16, ..., je nachdem das erste Bild bei dem Wurf mit der Nummer 1, 2, 3, 4, ... erscheint. Wie ist der Einsatz x zu bemessen?

Wir berechnen für A die Hoffnung h nach dem Zahlungsansatze. Als negative Zahlung tritt $z = -x$ auf mit $\mathfrak{W}(z) = 1$. Ferner erscheinen die positiven Zahlungen $2^1, 2^2, 2^3, \ldots$ mit den zugehörigen $\mathfrak{W}\mathfrak{W}$. $2^{-1}, 2^{-2}, 2^{-3}, \ldots$, woraus, da das Spiel auf n Würfe verabredet ist,

$$h = -x + 2^1(2^{-1}) + 2^2(2^{-2}) + \cdots + 2^n(2^{-n}) \tag{3}$$

folgt. Die Bedingung $h = 0$ liefert also für den Einsatz x den Wert n. Dieses Ergebnis hat nun viel Kopfzerbrechens verursacht, und zwar aus folgendem Grunde. Wenn n groß genommen wird, so ist die $\mathfrak{W}$., daß A auch nur seinen Einsatz wieder gewinnt, für das einzelne Spiel recht klein, weil die weitaus überwiegende Mehrzahl der möglichen Spielverläufe für A ungünstig steht. Beispielsweise müßte für $n = 64$ fünfmal hintereinander Schrift kommen, damit A wenigstens den Einsatz ausgezahlt erhält; die $\mathfrak{W}$. für ein solches Ereignis ist aber nur 1 : 32. Mit Rücksicht hierauf wird in den Darstellungen der W.-R. gesagt, daß im vorliegenden Falle die Regel von der mathematischen Hoffnung mit den Eingebungen des gesunden Menschen-

verstandes im Widerspruche stehe: jene lasse das Spiel auch für ein großes n als zulässig erscheinen, während dieser davon abrate. Infolgedessen hat man sich denn auch bemüht, für die Berechnung der Einsätze an Stelle der mathematischen Hoffnung andere Vorschriften ausfindig zu machen, bei denen man solche Widersprüche nicht zu fürchten habe. Es ist nicht zu leugnen, daß Aufgaben, wie das Petersburger Problem, für die hergebrachte Behandlungsweise in der Tat Schwierigkeiten bieten, allerdings nur deshalb, weil dabei Umstände übersehen werden, deren Bedeutung erst durch *v. Kries* (a. a. O. Seite 185 flg.) in bündiger Weise dargelegt worden ist. Bevor wir jedoch auf diesen Punkt eingehen, möge noch ein anderer Fall besprochen werden.

§ 47. Glücksspiele gelten im allgemeinen als verwerflich, sobald sie wegen der Höhe der Einsätze und der möglichen Verluste den Charakter einer harmlosen Unterhaltung verlieren. Diese Auffassung findet ihren bezeichnenden Ausdruck in der Tatsache, daß verschiedene Gesetzgebungen das gewerbsmäßige Hasardspiel strafrechtlich ahnden. Es gibt indessen auch Veranstaltungen, die ihrer Natur nach tatsächlich Zufallsspiele mit oft recht ansehnlichen Einsätzen sind, trotzdem aber als eine wohltätige und für das heutige Wirtschaftsleben unentbehrliche Einrichtung gelten, nämlich die Versicherungen. Das tritt am deutlichsten bei der eigentlichen Schadenversicherung hervor: der Versicherer ist der Spielhalter, der Versicherte der Einzelspieler, die sogenannte Prämie ist der Einsatz, und als Gewinn erscheint die Vergütung eines Schadens, dessen Eintreten vom Zufall abhängt und im allgemeinen nur mit geringer W. zu erwarten ist. Diese geringe W. bewirkt, daß ein großer Teil der Versicherten den Einsatz tatsächlich verliert. Ähnlich liegt die Sache bei der Lebensversicherung, nur daß hier bei dem Vorgange, auf den sich die Versicherung bezieht, zwar nicht das Eintreten an sich, wohl aber der Zeitpunkt des Eintretens dem Zufall unterworfen ist.

Die Ähnlichkeit mit der Spielbank reicht aber noch weiter. Große Spielbanken, wie sie früher in deutschen Bädern bestanden und anderwärts auch heute noch bestehen, rechnen damit, daß ihr Hauptgewinn aus den Verlusten zahlreicher Spieler mit kleinen und mittleren Einsätzen entspringt, denn nur wenn sehr viele einzelne Spiele stattfinden, ist diejenige Ausgleichung des Zufalls zu erwarten, welche bei der Ansetzung der Spielbedingungen zugrunde gelegt wird. Vereinzelte hohe Einsätze können ja Gewinn, gelegentlich aber auch, wenn es der Zufall so fügt, empfindliche Verluste bringen. Das Gleiche gilt von der Versicherung. Bezeichnet man als Risiko das in Geld ausgedrückte Interesse an dem Objekt, das der Versicherung unterworfen wird, so ist es unter sonst gleichen Umständen für die Stetigkeit des Versicherungsbetriebes weit wichtiger, daß recht viele, als

daß sehr hohe Risiken vorhanden sind. Dies wird noch deutlicher, wenn man den Ansatz betrachtet, der als Ausgangspunkt für die Berechnung der Prämie dient. Ist x die Prämie, g der versicherte Betrag oder das Risiko, $W(g)$ die W. seines Fälligwerdens, so erhält man, weil die Prämie der Regel nach dem Versicherer verbleibt, für die totale Hoffnung des Versicherers den Ausdruck

$$h = x \cdot 1 - g \cdot W(g). \tag{4}$$

Dieser Ansatz, der durch die Bedingung $h = 0$ aus g und $W(g)$ die Prämie bestimmt, beruht aber darauf, daß durch eine hinreichend häufige Wiederholung des vorliegenden Spieles eine genügende Ausgleichung der Zufälligkeiten erfolge, die dem Eintreten von g anhaften.

Das mit $h = 0$ berechnete x liefert die eigentliche Risikoprämie. Zu dieser treten als Zuschläge hinzu die anteiligen Beträge der Geschäftsunkosten und des in Aussicht genommenen Unternehmergewinns. Überdies darf man vermuten, daß in der Regel die in die Rechnung eingestellten $W(g)$ aus Gründen der Vorsicht eine merkliche Abrundung nach oben erfahren haben. Daraus folgt, daß der Versicherer mit einem positiven h arbeitet, daß also, an der mathematischen Hoffnung gemessen, das von dem Versicherten eingegangene Spiel für diesen nachteilig ist. Trotzdem gilt das Versicherungnehmen als eine vernünftige Maßregel, im Gegensatze zu dem Petersburger Problem, wo ein mit $h = 0$ berechneter Einsatz als unvernünftig gilt.

Die vorstehende Bemerkung lehrt, daß der Widerspruch zwischen den Vorschriften der Zufallstheorie und den Eingebungen des gesunden Menschenverstandes, den man bei dem Petersburger Problem finden zu müssen glaubte, in derselben Stärke, wenn auch mit umgekehrter Richtung, bei der Versicherung anzutreffen ist. Es ist sehr merkwürdig, daß man diese Tatsache so lange Zeit hindurch vollständig übersehen hat; andernfalls wäre man vermutlich schon früher dazu gelangt, dem Grunde dieser Widersprüche nachzugehen und zu erkennen, daß sie in Wirklichkeit nicht vorhanden sind, sondern lediglich auf einem falschen Gebrauche der als mathematische Hoffnung bezeichneten Größe beruhen. Um dies deutlich zu machen, sollen zunächst einige Bemerkungen über die sogenannten *Wertgleichungen* vorausgeschickt werden.

§ 48. Wenn zwei Dinge P und Q einander gleich gesetzt werden, so hat die Beziehung $P = Q$ je nach den Umständen einen sehr verschiedenen Sinn. Völlige Gleichheit ist imgrunde genommen nur dann vorhanden, wenn P und Q identisch sind, d. h. in allen an ihnen nachweisbaren Merkmalen übereinstimmen. Lassen wir diesen Fall beiseite, so wird sich die Gleichheit immer nur auf einen Teil der vorhandenen Merkmale beziehen können, während bezüglich der

übrigen Verschiedenheit besteht. Wird also für einen bestimmten Fall die Beziehung $P = Q$ angesetzt, so besagt dies, daß für den vorgelegten Fall nur die übereinstimmenden Merkmale von P und Q in Betracht kommen, und daß deswegen P und Q ohne weiteres miteinander vertauscht werden können. Dagegen hört die Vertauschung auf zulässig zu sein, sobald die weitere Behandlung des betrachteten Falles die Berücksichtigung noch anderer, nicht übereinstimmender Merkmale verlangt.

Im täglichen Leben hat man es nun fortwährend mit Gleichungen $P = Q$ zu tun, in denen die Bestandteile P und Q Dinge bedeuten, die so heterogen sind, wie z. B. eine Geldsumme und eine Tonne Korn oder eine Wattstunde elektrischer Energie oder das geistige Erzeugnis eines Schriftstellers. Hier ist das einzige übereinstimmende Merkmal durch den Wert gegeben, der jenen Dingen zuerkannt wird. Um diesen Umstand hervorzuheben, will ich kurz von der *Wertgleichung* $P = Q$ sprechen. Dies vorausgeschickt wollen wir die Frage stellen, ob eine Wertgleichung $P = Q$ auch noch bestehen bleibe, wenn links und rechts das Doppelte, das Dreifache usw. oder die Hälfte, das Drittel usw. von P und Q gesetzt wird. Das eine derartige *Erhaltung einer Wertgleichung*, wenigstens innerhalb gewisser Grenzen, sehr häufig stattfinde, wird man ohne Zweifel bejahen können. Wenn das Kilogramm einer Ware 1 Mark kostet, so werden für 10 Kilogramm 10 Mark gefordert und bezahlt. Allgemeiner gesprochen wird man die Erhaltung einer Wertgleichung als Regel da antreffen, wo es sich um Dinge handelt, die ständig einen bestimmten, wenn auch veränderlichen, Marktwert besitzen. Daraus darf aber nicht geschlossen werden, daß die Erhaltung der Wertgleichungen eine unverbrüchliche Regel sei, denn das Gegenteil ist ebenfalls eine durchaus häufige Erscheinung. Das goldene Zehnmarkstück und zehn silberne, als Scheidemünze ausgeprägte Einmarkstücke sind innerhalb des Geltungsbereiches der Markwährung gesetzlich einander gleich gestellt, auch wird ein ausländischer Bankier keine besondren Schwierigkeiten erheben, diese Gleichheit anzuerkennen, solange es sich nur um 10 Mark handelt. Soll dagegen eine Zahlung von 10000 Mark geleistet werden, so tritt im internationalen Geldverkehr sicher keine Erhaltung jener Wertgleichung ein, ja sie ist nicht einmal innerhalb des deutschen Reiches gesetzlich vorgeschrieben.

In dem vorstehenden Falle wird eine Wertgleichung durch Wiederholung zerstört; das Gleiche kann aber auch durch Teilung erfolgen. Wenn jemand für zwei benachbarte Hausgrundstücke, die je 50000 Mark wert sind, 100000 Mark anlegt, weil er die beiden Häuser zusammen für seinen Zweck gebrauchen kann, so folgt daraus keineswegs, daß diese Wertgleichung auch noch bei der Halbierung, d. h. wenn nur eines der beiden Häuser verkäuflich ist, bestehen bleiben müsse, denn

es ist sehr wohl möglich, daß das einzelne Haus für die Zwecke des Kauflustigen unbrauchbar ist.

§ 49. Wir wollen nun die vorstehenden Überlegungen auf die Zufallsspiele anwenden und zunächst den Fall des Massenspiels, d. h. der n-maligen Ausführung eines und desselben Spiels betrachten. Hierbei werden wir der Einfachheit halber die für unseren Zweck ausreichende Annahme machen, daß der Spieler A gegen einen festen Einsatz S im Falle des Gewinnens die seinen Einsatz einschließende Summe G ausgezahlt erhalte, während beim Verlieren der Einsatz verfällt. Die zu G gehörende W. bezeichnen wir mit $W(G)$ und die entsprechende partiale mathematische Hoffnung oder das Produkt aus G und $W(G)$ mit H. Dann setzt A bei dem einzelnen Spiele jedesmal einen sicheren Besitz S gegen ein vorläufig unsicheres Ding U, das je nach dem Verlaufe des Spieles den Wert Null oder G annimmt. Bezeichnet man noch die Gesamtheit der in Betracht kommenden S und U kurz mit $\sum S$ und $\sum U$, so besagt die Ausführung des auf n Einzelspiele ausgedehnten Massenspiels, daß A und sein Gegenspieler B die Wertgleichung $\sum S = \sum U$ für annehmbar erachten. Diese Auffassung, die vorläufig nur die beiden Spieler angeht, kann nun zutreffend oder verkehrt sein, so daß die Frage entsteht, wie jene Gleichung zu beurteilen sei.

Die linke Seite der Gleichung ist eine völlig bestimmte Größe, nämlich der n-fache Einsatz oder nS. Die rechte Seite $\sum U$ ist eine Vereinigung von ungewissen Dingen, erscheint also zunächst ebenfalls als etwas Ungewisses, das gleich irgend einem Vielfachen von G zwischen Null und nG sein kann. Diese Ungewißheit erfährt jedoch eine wesentliche Einschränkung, wenn wir, wie es jetzt geschehen soll, n als so groß voraussetzen, daß sich innerhalb des Massenspiels die den Einzelspielen anhaftenden Zufälligkeiten bis auf einen geringen Rest ausgleichen, daß also die tatsächliche rH. der gezahlten Gewinne bis auf einen unerheblichen Fehler gleich $W(G)$ wird. In diesem Falle darf $\sum U$ durch nH ersetzt werden, so daß das Urteil über die Wertgleichung von dem Betrage der Größe

$$nh = \sum U - \sum S = nH - nS \tag{5}$$

abhängt. Hierin steht jetzt eine sichere und genau bekannte Zahlung, nämlich nS, gegen eine andere sichere Zahlung, deren tatsächlicher Betrag zwar nicht im voraus genau bekannt ist, wohl aber mit ausreichender Annäherung durch den theoretischen Wert nH gegeben wird. Daher besitzt auch die Größe nh, die als der genäherte Reingewinn von A bezeichnet werden kann, eine reale Bedeutung. Bildet man ferner durch Division mit n die Gleichung

$$h = H - S, \tag{6}$$

so ist h der im *Durchschnitt* auf das einzelne Spiel fallende Anteil des Reingewinns. Hierbei ist jedoch zu beachten, daß vor der Hand h nur eine zur Bequemlichkeit der Rechnung eingeführte Hilfsgröße ist und für das einzelne Spiel lediglich insofern eine Bedeutung besitzt, als dieses Bestandteil eines Massenspiels ist. Mit dieser Einschränkung wird man kein Bedenken tragen, in der Größe h ein brauchbares Maß für den Vorteil oder Nachteil des Spielers A zu sehen.

§ 50. Die betrachtete Wertgleichung führt, wie man sieht, zu dem gewöhnlichen Hoffnungsansatz, sobald n hinreichend groß ist. Wesentlich anders stellt sich dagegen das Ergebnis, wenn nur ein isoliertes Spiel oder der Fall $n = 1$ in Frage kommt. Es handelt sich dann um die Wertgleichung $S = U$, die bei der herkömmlichen Betrachtungsweise durch die Substitution $U = H$ in $S = H$ verwandelt wird. Da nun die Zeichen S, U und H völlig disparate Dinge bedeuten, so ist es nach dem, was oben über die Teilung von Wertgleichungen bemerkt wurde, schlechterdings nicht erlaubt, die Substitution $U = H$ etwa dadurch rechtfertigen zu wollen, daß für ein hinreichend großes n die Gesamtheit $\sum U$ gleich nH gesetzt werden darf, vielmehr muß der Beweis für die Beziehung $U = H$, falls sie überhaupt zutrifft, aus den besonderen Umständen des isolierten Spiels hergeleitet werden. Um besser hervortreten zu lassen, was das besagen will, mögen bestimmte Zahlenwerte eingeführt werden. Es sei z. B. S gleich 1 Mark, G gleich 1000 Mark, $W(G)$ gleich 0.001, so daß H gleich 1 Mark wird. Dann soll laut der Substitution $U = H$ für die ungewisse Größe U, die *entweder* nichts *oder* 1000 Mark bedeutet, in dem isolierten Spiel unter allen Umständen der *bestimmte* Betrag von 1 Mark gesetzt werden dürfen. Das ist aber unter allen Umständen weiter nichts, als pure Willkür, denn wer auf ein isoliertes Spiel eingeht, der stellt seine Sache auf den Zufall: *entweder* er gewinnt *oder* er verliert, ein Mittleres dazwischen gibt es nicht — darüber hilft kein Grundsatz oder Lehrsatz der W.-R. hinweg. Des weiteren mutet die Gleichung $S = H$ dem Spieler zu, er müsse auf Grund der mathematischen Zufallstheorie die eine Mark, die er als sicheren Besitz in der Hand hält, gleich dem tausendsten Teile einer Summe achten, die allerdings 1000 Mark beträgt, die aber nur bei einem gar nicht vorherzusehenden Zusammentreffen von Umständen zur Auszahlung gelangt. Das ist ebenfalls bloße Willkür. Hieraus haben wir folgendes zu schließen: *Wenn in einem gegebenen Falle für ein einzelnes Spiel die Wertgleichung $S = U$ von einem Spieler als annehmbar erachtet wird, so sind die für ein solches Urteil maßgebenden Gründe gar nicht Sache der Zufallstheorie, sondern des Temperaments.* Der Eine kauft ein Lotterielos, weil er in seinem Herzen denkt, „weshalb sollen die schönen Treffer gerade bei meinem Lose vorbeigehen“; der Andere behält sein Geld in der Tasche, weil er sich nach dem Sprich-

wort richtet, daß der Sperling in der Hand besser sei, als die Taube auf dem Dache.

Die hier erörterten Verhältnisse liefern ein lehrreiches Beispiel dafür, wie unter Umständen der einem Dinge gegebene Name die Wirkung hat, daß dem Dinge Eigenschaften beigelegt werden, die mit seinen erkennbaren Merkmalen nicht notwendig zusammenhängen. Wenn das Produkt aus Gewinn und Wahrscheinlichkeit als mathematische Hoffnung bezeichnet wird, so ist das zunächst weiter nichts, als ein Name für eben dieses Produkt. Weil nun aber das Ding „Hoffnung" heißt, und weil es zugleich als eine in Mark und Pfennigen ausgedrückte Geldsumme auftritt, so schiebt sich in der hergebrachten Darstellung unvermerkt die Auslegung unter, daß einer solchen Summe ohne weiteres der gleiche Wert innewohne, wie einem sicheren physischen Besitz, der auf dieselbe Anzahl von Mark und Pfennigen lautet. Ich möchte die Vermutung wagen, daß diese Unterschiebung weniger leicht eingetreten wäre, wenn man statt der Bezeichnung „mathematische Hoffnung" irgend ein an sich nichtssagendes Wort, wie z. B. Abrakadabra oder dergleichen gewählt hätte.

Man kann nun noch fragen, wie sich die Sache gestaltet, wenn n zwar größer als Eins, aber nicht so groß ist, das eine hinreichende Ausgleichung des Zufalls eintritt. Die Antwort ergibt sich aus folgender Überlegung. In dem betrachteten Beispiel wird für die totale Hoffnung h der Ausdruck $H = S$ gesetzt, und zwar auf Grund der Substitution

$$H = \sum U : n. \tag{7}$$

Diese Substitution ist mit einem Fehler behaftet, dessen Spielraum von n abhängt. Ist n hinreichend groß, so ist der Fehler belanglos; läßt man dagegen n abnehmen, so wächst der zu befürchtende Fehler und erreicht, schon bevor n den niedrigsten Wert Eins annimmt, solche Beträge, daß die Substitution nicht einmal mehr als Annäherung brauchbar ist.

§ 51. Es ist vielleicht nicht überflüssig, im Anschlusse an die vorstehende Erörterung darauf hinzuweisen, wie der Hoffnungsansatz überall da beurteilt wird, wo ein Spiel nicht als harmloser Zeitvertreib oder zur Befriedigung einer gefährlichen Leidenschaft, sondern aus nüchterner Berechnung unternommen wird. Die Spielbank schreibt für die Einsätze ein Maximum vor und weist Beträge, die darüber hinausgehen, einfach zurück. Denn sie betrachtet Fälle, in denen ein einzelner wagehalsiger Spieler besonders hohe Summen setzen will, als Einzelspiele, die jedenfalls nicht oft wiederholt werden können, und die deshalb aus dem Rahmen des Massenspiels, auf dem die Spielbank aufgebaut ist, herausfallen. Jene Zurückweisung erfolgt überdies trotz des Umstandes, daß die Spielbank mit einem positiven Hoffnungswerte arbeitet, also nach dem Hoffnungsansatze von vorn-

herein günstiger gestellt ist als der Einzelspieler.[1]) In ähnlicher Weise und aus den gleichen Gründen werden von vorsichtigen Versicherungsanstalten die übermäßig hohen Risiken entweder einfach abgelehnt oder aber durch das Mittel der Rückversicherung teilweise abgewälzt.

Das Gegenstück hierzu liefert die Stellung des Versicherungnehmers: er spielt ebenfalls aus nüchterner Überlegung und trotz des Umstandes, daß sein Hoffnungswert negativ ist, denn er erachtet, ohne sich an die Hoffnungsregel zu kehren, den Nachteil aus einem möglichen, wenn auch wenig wahrscheinlichen Schaden für größer, als den sicheren physischen Verlust, den ihm die regelmäßige Zahlung der Prämien auferlegt. Dagegen tritt die Hoffnungsregel sofort in ihr Recht, wenn eine große Menge von Risiken in der Hand eines und desselben Besitzers vereinigt ist. Wollte z. B. eine große Eisenbahngesellschaft ihren Besitz, der seiner Hauptmasse nach ständig über weite Strecken zerstreut ist, in derselben Weise wie ein Privatmann versichern, so würde sie das negative Vorzeichen ihres Hoffnungswertes sehr bald fühlen; darum wird in solchen Fällen die sogenannte Selbstversicherung[2]) herangezogen.

§ 52. Faßt man das Gesagte kurz zusammen, so kann man sagen: der Hoffnungswert hat seinen guten Sinn bei einem Massenspiel, und zwar als eine Durchschnittsgröße, die zur Erleichterung der Rechnung eingeführt wird; er verliert dagegen seine Bedeutung bei dem Einzelspiel, weil bei diesem die Berechnung eines Durchschnitts sinnlos wird. Damit können wir nunmehr auch das Petersburger Problem rasch erledigen. Auf der einen Seite wird bei diesem Problem der Einsatz nach der Hoffnungsregel berechnet, die nur für ein Massenspiel Bedeutung hat. Auf der andern Seite sind die Spielbedingungen in

1) Ein lehrreiches und zugleich ergötzliches Geschichtchen berichtet *Em. Herrmann* in seinem Buche *„Die Theorie der Versicherung vom wirtschaftlichen Standpunkte. Wien, C. Konegen, 1897“* (Seite 29). Im Jahre 1853 gewann der an der Bank in Homburg spielende Charles Lucian Bonaparte, Fürst von Canino, in zwei Tagen die Summe von 200000 Gulden, weil er stets mit dem Maximum spielte. Am Abend des zweiten Tages beschlich die Bank die Ahnung, daß sie am andern Tage gesprengt werden könnte. Als Bonaparte am Morgen dieses dritten Tages wieder an der Bank erschien, um sein Spiel fortzusetzen, wurde er durch eine landgräfliche Verordnuug überrascht, welche in mehreren Abschriften an den Wänden des Spielsaales affichiert war und besagte, daß das Maximum der Spielbank auf 2000 Gulden herabgesetzt worden sei. Angesichts dieses Aktes der Spielbanksouveränität verließ Charles Lucian Bonaparte sofort den Spielsaal. Bald nach seiner Abreise wurde das Maximum laut landgräflicher Verordnung wieder auf 4000 Gulden erhöht.

2) Von manchen Schriftstellern wird die Selbstversicherung nicht als eigentliche Versicherung angesehen, weil dazu eine größere Anzahl von beteiligten Personen erforderlich sei. In Wahrheit kommt es jedoch bei dem Versicherungsspiel nicht auf die Menge der Personen, sondern auf die Menge derjenigen Risiken an, die im Sinne der Zufallstheorie als voneinander unabhängig anzusehen sind.

sinnreicher Weise so geregelt, daß eine für die Ausgleichung des Zufalls hinreichende Ausdehnung des Spiels nicht ernstlich in Frage kommen kann, sobald die Zahl der verabredeten Würfe einen irgendwie größeren Wert annimmt. Wenn z. B. das Spiel auch nur auf 10 Würfe verabredet wird, so ist die 𝔚. des höchsten Gewinnes 2^{10} durch den Wert

$$p = 1 : 1024$$

gegeben. Die 𝔚. für das Nichteintreten des Gewinnes ist dann bei dem einzelnen Spiel gleich $1 - p$ und bei einer Reihe von m Spielen gleich $(1 - p)^m$. Folglich wird die 𝔚., daß bei m Spielen der betrachtete Gewinn wenigstens einmal vorkommt, durch

$$\mathfrak{W} = 1 - (1 - p)^m$$

ausgedrückt. Der Verlauf dieser Größe ist am raschesten aus einem Täfelchen, wie dem folgenden, zu entnehmen.

$m = 500$	600	700	800	900	1000
$\mathfrak{W} = 0.39$	0.44	0.49	0.54	0.58	0.62.

Es sind also z. B. rund 700 Spiele nötig, damit der Spieler A den höchsten Gewinn wenigstens einmal mit $\mathfrak{W} = 0.5$ erwarten kann. Daraus ergibt sich zur Genüge, daß der Widerspruch, den man bei der betrachteten Aufgabe finden zu müssen glaubte, nur von einer verkehrten Anwendung der Hoffnungsregel herrührt, also in Wirklichlichkeit nicht vorhanden ist.

Die hier gegebene Klarlegung des besondern Verhaltens der Wertgleichungen und — im Zusammenhange damit — des Petersburger Problems hat ziemlich lange auf sich warten lassen und ist meines Wissens zuerst durch *J. von Kries* gegeben worden. Vorher weist die Literatur, namentlich die des achtzehnten Jahrhunderts, mancherlei zur Lösung des vermeintlichen Widerspruchs unternommene Versuche auf, von denen der bekannteste, nämlich die *moralische* Lösung von *Daniel Bernoulli*, hier noch der Vollständigkeit halber besprochen werden möge. Ich will dabei der Darstellung folgen, die *Laplace* im Schlußkapitel seiner „Théorie analytique . . .“ gegeben hat.

§ 53. Ein bestimmtes in Mark und Pfennigen ausgedrücktes, physisches Vermögen x hat für seinen Besitzer einen gewissen subjektiven Wert, der als das entsprechende *moralische Vermögen* y bezeichnet werden soll. Um die Beziehung zwischen x und y zu finden, wird festgesetzt, daß der unendlich kleinen Änderung dx des Vermögens x ein moralischer Wert dy zukommen solle, der einerseits proportional dx, andererseits umgekehrt proportional zu x ist, so daß

$$dy = K dx : x$$

wird, wo K eine positive Größe bedeutet, die von den besonderen

Umständen des vorliegendes Falles abhängt, von x aber unabhängig ist. Durch Integration folgt hieraus

$$y = K \log x + L, \tag{8}$$

wo L in demselben Sinne wie K eine Konstante ist.

Der vorstehende Ansatz ist offenbar nicht frei von Willkür, indessen kann man ihn vorläufig als „Arbeitshypothese" zulassen, weil er in einfacher Weise darauf Rücksicht nimmt, daß der subjektive Wert von dx um so kleiner ausfallen wird, je größer das bereits vorhandene Vermögen ist. Auch kann man sich auf die Analogie mit anderen Gebieten berufen: das *Weber*sche Gesetz für die Sinnesempfindungen mißt z. B. die Stärke einer Reizempfindung durch den Logarithmus der Intensität des entsprechenden physikalischen Reizes, und man könnte die Größe y ganz gut als den *empfundenen* Wert von x bezeichnen.

Die Konstante L kann durch passende Wahl der Maßeinheit von x gleich Null gemacht werden: wir wollen sie deshalb weiterhin einfach fortlassen. Die festgesetzt betrachten wir ein Glückspiel, bei dem der Spieler A die Änderungen $a, a', \ldots$ seines anfänglichen Vermögens v zu erwarten hat. Die 𝔚𝔚. dieser Änderungen seien $W(a), W(a'), \ldots$, also die Summe der $W(a)$ gleich Eins. Dann kommen als mögliche Endergebnisse des Spiels die Vermögenswerte

$$y = K \log (v + a), \qquad y' = K \log (v + a'), \ldots$$

in Betracht, deren Erwartungswerte durch die Produkte

$$y W(a), \qquad y' W(a'), \ldots$$

gemessen werden, womit man für die totale Erwartung den Ausdruck

$$Y = y W(a) + y' W(a') + \cdots$$

erhält. Das physische Vermögen X, das dem moralischen Werte Y entspricht, ist durch

$$Y = K \log X$$

gegeben, so daß

$$\log X = W(a) \log (v + a) + W(a') \log (v + a') + \cdots \tag{9}$$

wird. Die Differenz $X - v$ liefert dann die *moralische Hoffnung*, die bei der vorliegenden Betrachtungsweise an die Stelle der früher benutzten mathematischen Hoffnung tritt.

Die Größe Y hat, wenn man sie genauer ansieht, offenbar folgende Bedeutung. Man denke sich, daß A das Spiel unbeschränkt oft ausführt, und zwar jedesmal mit demselben Anfangsvermögen v beginnend, dann ist Y das arithmetische Mittel oder der Durchschnitt aus allen y, mit denen die einzelnen Spiele abschließen. Hieraus folgt, daß sich bei einem isolierten Spiel gegen den Gebrauch von Y genau dieselben

Einwände erheben lassen, wie oben gegen die mathematische Hoffnung. Handelt es sich dagegen um ein Massenspiel, so müßte man bei der Berechnung von Y die Gesamtheit der einzelnen Spiele als ein einziges Spiel auffassen, da ja das Anfangsvermögen, das jedem neuen Spiel zugrunde zu legen ist, von dem Ausfall der vorhergegangenen Spiele abhängt. Wir wollen indessen bei diesen Einwänden nicht verweilen, sondern vorerst zusehen, wie sich jetzt die Lösung des Petersburger Problems gestaltet.

§ 54. Nimmt A das auf n Würfe verabredete Spiel bei dem Anfangsvermögen v mit dem noch festzustellenden Einsatze x an, und soll ihm ferner das Spiel, moralisch gemessen, weder Vorteil noch Nachteil bringen, so muß in (9) X gleich v werden. Weiter sind, wenn das erste Bild beim p-ten Wurf erscheint, das geänderte Vermögen und die zugehörige 𝔚. durch die Ausdrücke

$$v - x + 2^p \quad \text{und} \quad 2^{-p}$$

gegeben, woraus in (9) der Term

$$2^{-p} \log (v - x + 2^p)$$

entsteht. Dazu kommt dann noch der Fall, daß immer nur Schrift geworfen wird. Dieser liefert das geänderte Vermögen $v - x$ mit 𝔚 $= 2^{-n}$, woraus in (9) der Term

$$2^{-n} \log (v - x)$$

entspringt. Faßt man nun zusammen und substituiert für X das Anfangsvermögen v, so erhält man damit zur Bestimmung von x die Bedingung

$$\log v = 2^{-n} \log (v - x) + \sum_p 2^{-p} \log (v - x + 2^p), \quad (p = 1, 2, \ldots n). \quad (10)$$

Es ist für unseren Zweck nicht nötig, nach der Auflösung dieser Gleichung zu suchen, sondern es wird genügen, den Einfluß der höheren Gewinne abzuschätzen.

Da man in (10) die natürlichen Logarithmen ohne weiteres durch die gemeinen ersetzen kann, die mit Log bezeichnet werden sollen, so kann man, wenn

$$v = w + x$$

gesetzt wird, schreiben

$$\operatorname{Log} (w + x) = 2^{-n} \operatorname{Log} w + \sum_p 2^{-p} \operatorname{Log} (w + 2^p). \quad (11)$$

Zieht man hiervon die mit $\operatorname{Log} w$ multiplizierte Gleichung

$$1 = 2^{-n} + \sum_p 2^{-p}$$

ab, so wird

$$\operatorname{Log} \frac{w + x}{w} = \sum_p 2^{-p} \operatorname{Log} \frac{w + 2^p}{w}, \quad (12)$$

worin p von 1 bis n zu laufen hat. Berechnet man nun auf der rechten Seite für ein gegebenes w, z. B. für $w = 10000$, die einzelnen Glieder der Reihe, so erkennt man, daß die höheren Glieder, die ja den höheren Gewinnen entsprechen, zu dem Summenwert nur wenig beitragen und folgeweise auch nur geringen Einfluß auf den Wert von x ausüben werden. Der Einsatz x wird also sehr nahe so gefunden, als ob die höheren Gewinne überhaupt nicht vorhanden wären. Gegen dieses Ergebnis ist natürlich nichts einzuwenden, indessen braucht man nicht erst den moralischen Ansatz heranzuziehen, um bei der Beurteilung des Einzelspiels die Aussichten des Spielers A von vornherein einigermaßen richtig zu würdigen. Wichtiger als dieser Umstand ist jedoch ein anderer Punkt, der jetzt noch besprochen werden soll.

Bildet man, zu dem allgemeinen Ausdruck (9) zurückkehrend, mit der auf positive Werte beschränkten Veränderlichen t den Ausdruck

$$Z(t) = W(a) \log(v + ta) + W(a') \log(v + ta') + \cdots,$$

so wird $Z(0) = \log v$, da ja die Summe der $W(a)$ gleich Eins ist. Ferner erhält man für $Z(1)$ den Wert $\log X$. Danach ist X größer oder kleiner als v, je nachdem $Z(1)$ größer oder kleiner als $Z(0)$ ist, d. h. das Vorzeichen der moralischen Hoffnung $X - v$ stimmt mit dem Vorzeichen der Differenz $Z(1) - Z(0)$ überein. Bildet man nun die zweite Ableitung von Z nach t, so ist diese beständig negativ, d. h. die erste Ableitung nimmt mit wachsendem t beständig ab. Für $t = 0$ nimmt die erste Ableitung mit der Abkürzung

$$h = a\,W(a) + a'\,W(a') + \cdots$$

den Wert $h : v$ an, wobei h offenbar gleich der mathematischen Hoffnung ist. Läßt man jetzt t von 0 bis 1 gehen, so beginnt die erste Ableitung mit dem Werte $h : v$, ist also in dem betrachteten Intervall durchweg negativ, sobald h Null oder negativ ist. In diesem Falle nimmt daher $Z(t)$ von Anfang an beständig ab, d. h. die moralische Hoffnung $X - v$ ist sicher negativ. Läßt man ferner h über Null hinauswachsen, so wird wegen der Stetigkeit das negative Vorzeichen von $X - v$ anfangs noch bestehen bleiben und einen Wechsel erst dann erfahren, wenn h einen gewissen positiven Wert erreicht hat. Daraus folgt, daß ein nach der Bedingung $h = 0$ geordnetes Spiel für alle Parteien moralisch nachteilig ist, und daß es ferner kein Spiel gibt, bei dem nicht wenigstens eine Partei moralischen Nachteil erleidet.

Der vorstehende Satz ist häufig mit einer gewissen Emphase betont worden, weil er den mathematischen Nachweis für das Unmoralische aller Glücksspiele erbringe. Dabei hat man jedoch vergessen, daß die Versicherung auch zu den Glücksspielen gehört, und daß sie

eine wirtschaftlich nützliche und oft genug auch sittlich wertvolle Einrichtung ist. Wenn man einmal von den beiden hier besprochenen Lösungen des Petersburger Problems die moralische für die vernünftigere hält, so ist diese Bewertung logischerweise auch auf die anderen Glücksspiele auszudehnen, und es wäre gerade bei der Versicherung eine ausgezeichnete Gelegenheit gegeben, die Vorzüge der moralischen Rechnung einer beweiskräftigen Ernstprobe zu unterwerfen. Ein solcher Versuch ist aber, da sein Ausgang vorherzusehen ist, wohlweislich niemals gemacht worden. *Es ist ja nicht undenkbar, daß man in der W.-R. einmal auf Probleme geführt wird, bei denen der moralische Ansatz angebracht ist; einstweilen wird man ihn jedoch, soweit zufällige Ereignisse in Betracht kommen, für eine Künstelei halten müssen, die zur Beseitigung eines vermeintlichen Widerspruchs ersonnen wurde, die aber bei der Anwendung auf ernsthafte Dinge nicht Stich hält.* Daß dagegen auf anderen Gebieten, also außerhalb der W.-R., die von *Daniel Bernoulli* benutzte Vorstellung eine Bedeutung besitzen kann, ist oben bereits kurz erwähnt worden.

Achte Vorlesung.

Geometrische Wahrscheinlichkeiten.

§ 55. Da die W.-R. ihrem Wesen nach eine Häufigkeitsrechnung ist, so sind ihre Methoden, wie schon früher bemerkt wurde, auch bei solchen Aufgaben anwendbar, bei denen es sich nicht um zufällige Ereignisse im gewöhnlichen Sinne dieses Wortes handelt. Eines der interessantesten Beispiele dafür dürfte die Anwendung der W.-R. in der kinetischen Gastheorie sein, deren Betrachtung uns allerdings über die hier gesteckten Grenzen hinausführen würde. Ich will mich deshalb darauf beschränken, einige Fälle aus der Klasse der sogenannten geometrischen WW. zu behandeln.

Sind in einer Ebene zwei Kurven A und B gegeben, so kann man fragen, welche W. dafür bestehe, daß eine blindlings gezogene Gerade, wenn sie A trifft, auch B treffe. Zur Lösung der Aufgabe hat man offenbar die Menge der Sekanten von A in eine Reihe von gleichberechtigten Gliedern zu ordnen, dann daraus die Menge der günstigen Fälle, d. h. der B treffenden Sekanten abzuspalten und schließlich den Quotienten der beiden Mengen zu bilden. Man erkennt hierbei sofort, daß das Ergebnis wesentlich von den Bedingungen abhängen wird, nach denen man die Gleichberechtigung der Sekanten von A beurteilt. Diese Bedingungen sind, wenn es sich, wie im vorliegenden Falle, um beliebig ersonnene Aufgaben handelt, innerhalb gewisser Grenzen der Willkür überlassen, so daß man jedesmal

bestimmt vorzuschreiben hat, wie die Gleichberechtigung bemessen werden soll. In dieser Beziehung wollen wir für die weiterhin behandelten Beispiele folgendes festsetzen.

Handelt es sich um die Punkte einer Linie, so soll die Menge der Punkte durch die Länge der Linie gemessen werden. Ebenso messen wir die Menge der Punkte eines Flächenstücks oder eines Körpers durch den Inhalt der Fläche oder des Körpers. Da hiernach gleichgroße Elemente dieser Gebilde jedesmal gleiche Punktmengen enthalten, so können wir auch sagen, daß bei der getroffenen Bestimmung die Verteilung der Punkte überall die gleiche Dichtigkeit besitze.

Um eine Menge von Geraden in einer Ebene abzuzählen, schreiben wir zunächst die Gleichung der Geraden in der Gestalt

$$x \cos f + y \sin f = a, \tag{1}$$

wo a das vom Nullpunkt der Koordinaten x, y auf die Gerade gefällte Lot und f die Richtung dieses Lotes bedeutet. Beschränken wir a auf positive Werte, so hat f, damit alle Geraden erhalten werden, den vollen Kreisumfang zu durchlaufen. Läßt man dagegen für a sowohl positive, wie negative Werte zu, so ist f auf einen Halbkreis zu beschränken. Die Größen a und f denken wir uns als rechtwinklige Koordinaten eines Ebenenpunktes abgetragen und messen die Menge der Geraden, deren Parameterwerte a, f zu allen Punkten eines gewissen Stückes der Ebene gehören, durch den Inhalt eben dieses Stückes. Hiernach wird ein Element der betrachteten Geradenmenge durch das Produkt $da \cdot df$ gemessen.

§ 56. Nach diesen Festsetzungen gehen wir jetzt zu einigen Aufgaben über, die sich auf Kurven in einer Ebene beziehen. Wir betrachten zunächst eine geschlossene Linie A, die ohne Verschlingungen verläuft und überall nach außen konvex ist. Dann wird mit Rücksicht auf (1) die Menge $M(A)$ der Geraden, die A treffen, durch die Summe der zu diesen Geraden gehörigen Elemente $da \cdot df$ gemessen. Bei der Berechnung dieser Summe beginnen wir mit der Summation nach a, betrachten also zuerst nur diejenigen Geraden G, die zu einem festen f gehören. Diese Geraden sind parallel, und die vom Nullpunkte zu ihnen gezogene Senkrechte S hat die Richtung f. Die Durchschnitte der betrachteten G mit der Senkrechten S bedecken auf der letzteren eine gewisse Strecke b, die nichts anderes als die gesuchte Summe der da ist. Die Größe a tritt hierbei als eine längs S gezählte Abszisse auf, die sowohl positiv, wie negativ sein kann. Ist P ein Punkt, der den Linienzug A einmal durchläuft, und projiziert man P auf S, so durchläuft diese Projektion die Strecke b zweimal, und zwar in entgegengesetzten Richtungen. Bedeutet nun ds ein Bogenelement von A, und g die Richtung von ds,

ist ferner $|x|$ das Zeichen für den absoluten Betrag der Größe x, so erhält man für den absoluten Betrag der Projektion von ds auf S den Ausdruck

$$|\cos(f-g)|\, ds.$$

Dieser Ausdruck, über alle ds summiert, liefert den doppelten Betrag von b, so daß

$$2\int da = 2b = \int |\cos(f-g)|\, ds$$

wird. Daraus folgt für das gesuchte $M(A)$ die Darstellung

$$2\, M(A) = 2\int da \cdot df = \int |\cos(f-g)|\, ds \cdot df,$$

in der nach f über einen Halbkreis zu integrieren ist. Nun wird, wenn man jetzt zuerst nach f integriert, und dabei f erst von g bis $g+\frac{1}{2}\pi$ und dann von $g+\frac{1}{2}\pi$ bis $g+\pi$ laufen läßt,

$$\int |\cos(f-g)|\, df = 2,$$

woraus die einfache Beziehung

$$M(A) = \int ds = \text{Umfang von } A$$

folgt.

Die vorstehende Herleitung versagt, wenn A Einbiegungen oder Schlingen besitzt, weil dann die Projektion von ds einzelne Stellen von b mehr als zweimal bedeckt. In diesem Falle denken wir uns um A einen Faden herumgelegt und straff angespannt, dessen Verlauf wir kurz die *Umspannung* von A nennen wollen. Bezeichnet $U(A)$ der Einfachheit halber zugleich die umspannende Linie und deren Länge, so ist, wie man leicht erkennt, jede Sekante von A auch Sekante von $U(A)$, und umgekehrt. Da nun für $U(A)$ die obige Beweisführung ohne weiteres bestehen bleibt, so wird jetzt

$$M(A) = M(U(A)) = U(A). \tag{2}$$

Diese Gleichung gilt auch noch für eine ungeschlossene Linie A, wie man sofort erkennt, wenn man sich A zweimal, und zwar in entgegengesetzten Richtungen, durchlaufen denkt und diese Punktbahn als eine ausgeartete geschlossene Linie ansieht. Im besonderen wird, wenn A eine geradlinige Strecke von der Länge L bedeutet, $M(A)$ gleich $2L$.

Denkt man sich für A einen unendlich kleinen Kreis K gewählt, den man nach und nach an beliebige Stellen der Ebene verlegt, so bleibt die Menge der Geraden, die K schneiden, immer dieselbe. Daraus folgt, daß die nach (1) geordnete Verteilung der Geraden überall die gleiche Dichtigkeit besitzt.

§ 57. Ist B eine Kurve, die innerhalb der geschlossenen Kurve A liegt, so wollen wir jetzt fragen, welche W. dafür bestehe, daß eine blindlings gezogene Sekante von A auch B treffe. Die Menge der

gleichberechtigten Fälle wird durch $U(A)$ gemessen, die Menge der günstigen Fälle durch $U(B)$, folglich ist die gesuchte W. gleich dem Quotienten

$$U(B) : U(A).$$

Aus der vorstehenden Formel lassen sich verschiedene Folgerungen ziehen. Eine horizontale Ebene sei mit einem Gitter von äquidistanten Parallelen bedeckt, die den Abstand a besitzen; auf dieses Gitter wird eine Scheibe B geworfen, die so gestaltet ist, daß sie nie mehr, als eine einzige Gittergerade treffen kann: gesucht wird die W., daß B eine der Geraden trifft. Zu dem Ende denken wir uns zunächst mit B einen umschließenden Kreis A vom Durchmesser a fest verbunden, dann wird A stets von einer der Geraden geschnitten. Sind nun die geworfenen Orte des Kreismittelpunktes und ebenso die geworfenen Richtungen eines im Kreise festen Durchmessers gleich dicht verteilt, so werden auch die Kreissehnen, die die Würfe aus den Gitterlinien ausschneiden, gleich dicht verteilt sein. Dann ist aber die gesuchte W. nichts anderes, als die W., daß eine Sekante des Kreises auch die Scheibe trifft, woraus

$$\mathfrak{W} = U(B) : U(A) = U(B) : a\pi$$

folgt.

Läßt man die Scheibe in ein unendlich schmales Rechteck von der Länge L übergehen, so wird

$$\mathfrak{W} = 2\,L : a\pi. \tag{3}$$

Will man die vorstehende Formel experimentell prüfen, so kann man als Rechteck eine dünne zylindrische Nadel benutzen und erhält damit einen besonderen Fall des *Buffon*schen *Nadelproblems*, dessen allgemeinere Gestalt darauf hinauskommt, daß man statt des Parallelengitters ein aus kongruenten Maschen gebildetes Netz aufzeichnet. Für die Formel (3) hat *R. Wolf* eine Probe angestellt, indem er mit einer Nadel von 36 mm Länge 5000 Würfe auf ein Parallelengitter von 45 mm Geradenabstand ausführte. Die beobachtete rH. der günstigen Fälle lieferte für $\mathfrak{W}$ den Wert 0.5064, während aus (3) der theoretische Wert 0.5093 folgt. Die Übereinstimmung zwischen den beiden Zahlen ist, worauf ich an dieser Stelle nicht näher eingehen will, als befriedigend anzusehen.[1])

§ 58. Die betrachtete Formel läßt noch eine andere Folgerung zu, die zu der *statistischen* Methode für die Rektifikation von Kurven

---

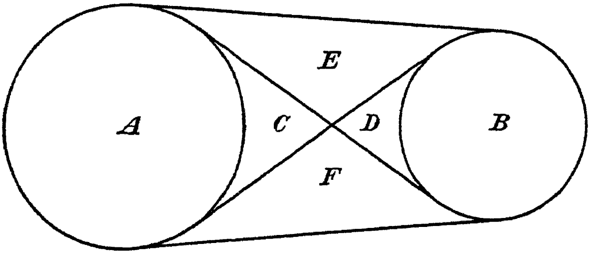
1) In dem Buche von *E. Czuber* „Geometrische Wahrscheinlichkeiten und Mittelwerte" (Leipzig, B. G. Teubner, 1884) ist eine sehr reichhaltige Zusammenstellung von Untersuchungen über geometrische Häufigkeiten gegeben. Ebendaselbst (Seite 84 flg.) findet man auch näheres über das Nadelproblem und die Versuche von *R. Wolf*.

führt. Der Quotient der beiden Umfänge $U(A)$ und $U(B)$ stellte sich als Quotient von zwei unendlichen Sekantenmengen dar, woraus zu schließen ist, daß man einen genäherten Wert des Quotienten erhalten wird, wenn man statt der stetig veränderlichen Sekanten eine gleichmäßig angeordnete, diskrete Menge zugrunde legt und an dieser die geforderten Auszählungen vornimmt. Soll auf diese Weise z. B. der Wert von $U(B)$ für eine gezeichnete Figur B gefunden werden, so kann man folgenden Weg einschlagen. Man zieht um B einen umschließenden Kreis A von passendem Durchmesser, legt auf die Zeichnung ein durchsichtiges Blatt mit einem Parallelengitter von z. B. 1 mm Geradenabstand und zählt bei dieser Lage des Blattes die Menge der Gittersekanten von A und B ab. Dann dreht man das Blatt um einen aliquoten Teil des Halbkreises, zählt wieder, und fährt mit dem Drehen und Zählen fort, bis man zu einer letzten Stellung des Blattes gelangt, von der aus die nächstfolgende Drehung wieder auf die Anfangsrichtung der Parallelen zurückführen würde. Nachdem diese Zählungen erledigt sind, erhält man den genäherten Quotienten von $U(B)$ durch $U(A)$, wenn man die gezählte Gesamtmenge der B-Sekanten durch die entsprechende Menge der A-Sekanten dividiert. Daraus folgt dann ohne Mühe der Wert von $U(B)$, weil ja der Wert von $U(A)$ unmittelbar aus dem Durchmesser von A zu finden ist.

Sieht man sich das beschriebene Verfahren genauer an, so erkennt man leicht, daß der Kreis A in Wirklichkeit gar nicht gezogen zu werden braucht, und daß man $U(B)$ findet, wenn man aus den für jede Blattlage gezählten Mengen der B-Sekanten das arithmetische Mittel bildet und letzteres mit π multipliziert. Fügt man ferner zu der für die einzelne Blattlage gezählten Menge jedesmal die, in Bruchteilen des Millimeters ausgedrückten, Pfeilhöhen der in der Regel übrigbleibenden äußersten Segmente von B hinzu, so hat das dieselbe Wirkung, als wenn das Gitter statt des Geradenabstandes von 1 mm einen solchen von 0.1 mm besäße. Dadurch wird offenbar eine Erhöhung der Genauigkeit der Abzählungen erreicht, während im übrigen der Grad der Annäherung von der Anzahl der benutzten Blattlagen abhängt.

§ 59. Wir betrachten jetzt zwei konvexe Figuren A und B, die einander ausschließen, und fragen nach der Menge der Geraden, die A und B gleichzeitig treffen. Zur Lösung der Aufgabe denken wir uns A und B gleichzeitig durch einen straffen Faden umspannt, was auf doppelte Art geschehen kann, nämlich einmal mit einer Kreuzung zwischen A und B, sodann aber auch ohne Kreuzung. Die Länge des ersten Fadens nennen wir X, die des zweiten Y und bezeichnen die zwischen A und B von den Fäden abgegrenzten Ebenenstücke mit C, D, E, F in der aus umstehender Figur ersichtlichen Weise.

Für die Sekantenmengen und die Umspannungen der sechs vorliegenden Ebenenstücke benutzen wir wie bisher die Zeichen $M(A), \ldots$ und $U(A), \ldots$, in der gleichen Weise gebrauchen wir die Zeichen $M(A + C), \ldots$ und $U(A + C), \ldots$, um die M und U der durch Zusammenlegung von A und C, usw. erhaltenen Ebenenstücke $A + C$, usw. anzugeben. Endlich soll das



Zeichen $M(P, Q)$ die Menge der Geraden bedeuten, die zwei Figuren P und Q gleichzeitig treffen.

Jede Sekante von A ist auch Sekante von $A + C$, und das gleiche gilt von den beiden Figuren B und $B + D$. Ebenso ist jede gemeinsame Sekante von A und B zugleich auch gemeinsame Sekante der Figuren $A + C$ und $B + D$. Betrachtet man ferner eine gemeinsame Sekante S von $A + C$ und $B + D$, so erkennt man aus der Figur, daß S auch gemeinsame Sekante von A und B ist, denn eine Gerade, die zwar C und D, nicht aber A und B trifft, kann nicht gezogen werden. Infolgedessen wird

$$M(A, B) = M(A + C, B + D), \tag{4}$$

so daß wir bei unserer Aufgabe an die Stelle von A und B die Stücke $A + C$ und $B + D$ setzen dürfen.

Die Sekantenmenge $M(A + C)$ läßt sich in zwei Teile spalten, je nachdem von der einzelnen Sekante auch $B + D$ getroffen wird oder nicht. Mit Rücksicht hierauf können wir die unmittelbar verständliche Gleichung

$$M(A + C) = M(A + C, B + D) + M(\text{nur } A + C)$$

ansetzen, der dann auf der anderen Seite des Kreuzungspunktes die Gleichung

$$M(B + D) = M(A + C, B + D) + M(\text{nur } B + D)$$

entspricht. Addiert man beide Gleichungen und beachtet dabei die Beziehungen

$$M(A + C) = U(A + C), \quad M(B + D) = U(B + D),$$
$$U(A + C) + U(B + D) = X,$$

so kann man schreiben

$$X = M(A + C, B + D) + Z, \tag{5}$$
$$Z = M(A + C, B + D) + M(\text{nur } A + C) + M(\text{nur } B + D).$$

Betrachtet man jetzt die Figur $P = A + C + B + D$, so besteht P offenbar aus zwei konvexen Blättern, die in einer Spitze zusammenhängen. Die Sekantenmenge $M(P)$ läßt sich in drei Bestandteile gliedern, je nachdem nur das erste Blatt oder nur das zweite oder aber beide getroffen werden. Danach ist Z nichts anderes als $M(P)$. Ferner ist $U(P)$ durch den Faden Y gegeben, also

$$Z = M(P) = U(P) = Y.$$

Damit liefert aber die Verbindung von (4) und (5) die einfache Gleichung

$$M(A, B) = X - Y. \tag{6}$$

§ 60. Des weiteren wollen wir nun noch nach der Menge N der Geraden fragen, die zwischen A und B frei hindurchgehen. Die Menge $M(P)$ oder Y enthält folgende Bestandteile: 1) die Menge N, 2) die Menge $M(A)$ unter Ausschluß von $M(A, B)$, 3) die Menge $M(B)$ unter Ausschluß von $M(A, B)$, 4) die Menge $M(A, B)$. Demgemäß setzen wir an:

$$Y = N + M(\text{nur } A) + M(\text{nur } B) + M(A, B),$$

ferner

$$M(A) = M(\text{nur } A) + M(A, B),$$
$$M(B) = M(\text{nur } B) + M(A, B),$$

woraus zunächst

$$\begin{aligned} Y &= N + M(A) + M(B) - M(A, B) \\ &= N + U(A) + U(B) - (X - Y) \end{aligned}$$

und schließlich

$$N = X - U(A) - U(B) \tag{7}$$

folgt.

Wenn die beiden Figuren A und B übereinander greifen, so betrachten wir zunächst das Ebenenstück Q, das von A und B, sei es allein, sei es gemeinsam, überdeckt wird, und setzen an:

$$\begin{aligned} M(Q) &= M(\text{nur } A) + M(\text{nur } B) + M(A, B) \\ &= [M(\text{nur } A) + M(A, B)] + [M(\text{nur } B) + M(A,B)] - M(A,B). \end{aligned}$$

Da nun $M(A)$ aus den beiden Mengen $M(A, B)$ und $M(\text{nur } A)$ besteht, und entsprechendes für $M(B)$ gilt, so wird

$$M(Q) = M(A) + M(B) - M(A, B),$$

woraus

$$M(A, B) = U(A) + U(B) - U(Q) \tag{8}$$

folgt.

§ 61. Im Anschlusse an die bisher behandelten Aufgaben untersuchen wir noch die Menge der Schnittpunkte zwischen den Sekanten einer konvexen Figur A, die den Umfang U und den Inhalt J besitzt. Die Geraden der Ebene lassen wir wieder durch die Gleichung

$$x \cos f + y \sin f = a$$

von den Parametern a und f abhängen und denken uns eine bestimmte Geradenmenge P dadurch abgegrenzt, daß wir die Geraden einem gewissen Bedingungssystem C unterwerfen. Dann ist P gleich der Summe der Elemente $da \cdot df$, die den Bedingungen C genügen. Des weiteren denken wir uns die Menge P wiederholt, bezeichnen sie jedoch jetzt zum Unterschiede mit Q und ihre Parameter durch b, g an Stelle von a, f. Das zu $da \cdot df$ gehörige Mengenelement P erzeugt mit den Geraden des Mengenelementes $db \cdot dg$ von Q eine Anzahl von Durchschnittspunkten, deren Menge wir durch das Produkt

$$(da \cdot df)(db \cdot dg) \tag{9}$$

messen. Um die Menge M der Durchschnitte zwischen den Geraden von P zu erhalten, suchen wir die Menge N derjenigen Durchschnitte auf, die die Geraden von P mit denen von Q erzeugen. Zu dem Ende ist das Produkt (9) über alle Parameterwerte zu summieren, die den vorgeschrieben Bedingungen C genügen. Hierbei tritt der Durchschnitt zwischen zwei beliebig herausgegriffenen Geraden F und G jedesmal doppelt auf, indem das eine Mal F zu P und G zu Q, das andere Mal F zu Q und G zu P gehören kann. Danach ist N gleich $2M$ und

$$2M = \int da\, df\, db\, dg, \tag{10}$$

wo die Integration über alle den Bedingungen C genügenden Parameterwerte auszuführen ist.

Als Bedingung C schreiben wir jetzt vor, daß die Geraden von P Sekanten von A sein sollen. Halten wir zunächst b und g fest, summieren also zuerst nach a und f, so wird die Summe der Elemente $da \cdot df$ gleich der Menge der Sekanten von A oder gleich dem Umfang U, folglich ist

$$2M = U \int db\, dg.$$

Hieraus ergibt sich, wenn nun nach b und g summiert wird,

$$2M = U^2. \tag{11}$$

Die Menge M zerfällt in zwei Bestandteile M_i und M_a, je nachdem die betrachteten Schnittpunkte innerhalb oder außerhalb A liegen. Um diese beiden Teilmengen zu finden, hat man wieder das Produkt (9) zu summieren, jedoch mit der engeren Bedingung, daß die Sekantenschnitte zu M_i oder M_a gehören müssen. Wir halten hierbei zunächst wieder b und g fest und bezeichnen die von der Geraden (b, g) erzeugte Sehne von A mit S. Dann kommen bei der Berechnung von M_i unter den Geraden (a, f) nur diejenigen in Betracht, die S treffen. Die Menge dieser Geraden ist gleich der Umspannung von S, also gleich $2S$, wenn wir unter S zugleich die Sehnenlänge verstehen. Damit wird

$$2M_i = \int 2S\, db\, dg.$$

Summiert man hierin bei festem g zuerst nach b, so ist die Summe der Produkte $S \cdot db$ gleich dem Inhalt J der Figur A, denn die db sind die Abstände zwischen den aufeinander folgenden S. Daraus folgt

$$M_i = J \int dg = \pi J, \tag{12}$$

und weiter

$$2 M_a = 2 M - 2 M_i = U^2 - 2 \pi J. \tag{13}$$

Die Menge M_i, dividiert durch die Fläche J liefert die mittlere Dichte der inneren Schnittpunkte. Diese ist nach (12) gleich π. Zeichnet man nun innerhalb A einen beliebig kleinen Kreis mit dem Inhalt K, so ist jede Sekante von K auch Sekante von A. Daraus folgt, daß von der Menge M_i ebensoviele Schnittpunkte in K fallen, als die Menge der inneren Sekantenschnitte der für sich betrachteten Figur K beträgt, d. h. innerhalb K liegt von der Menge M_i die Teilmenge πK. Damit wird die mittlere Dichte der in K fallenden Schnitte M_i wieder gleich π, also ist, wenn K unendlich klein gewählt wird, die Dichte der Schnitte M_i überall konstant und gleich π.

§ 62. Um die Schnittdichten für M_a zu erhalten, betrachten wir eine unendlich kleine geradlinige Strecke von der Länge S, die außerhalb A auf einer Sekante von A liegt. Die Sekantenmenge, die S trifft, wird nach (6) gleich der Differenz zwischen der gekreuzten und der ungekreuzten gemeinsamen Umspannung von A und S. Ist nun h der Winkel zwischen den beiden von S aus an A gezogenen Tangenten, sind ferner p und q die Winkel zwischen diesen Tangenten und S, so wird jene Differenz, wie man leicht aus einer dafür zu zeichnenden Figur erkennt, gleich

$$2 S - S \cos p - S \cos q.$$

Dies vorausgeschickt betrachten wir einen unendlich kleinen Kreis mit dem Inhalte K außerhalb A und fragen nach der Menge N der Schnitte, die von M_a innerhalb K fallen. Zu dem Ende ist in dem Ausdrucke

$$2 N = \int da\, df\, db\, dg$$

über alle Geraden (a, f) und (b, g) zu summieren, die gemeinsame Sekanten von A und K sind. Halten wir zunächst wiederum b und g fest, so wird auf der Geraden (b, g), die ja Sekante von A sein muß, von dem Kreise K eine gewisse Sehne S ausgeschnitten. Die Menge der S treffenden Sekanten von A haben wir bereits ermittelt und man erhält damit

$$2 N = \int (2 S - S \cos p - S \cos q)\, db\, dg,$$

wo die Winkel p und q die vorhin angegebene Bedeutung besitzen. Summiert man weiter nach b bei festem g, so bleiben hierbei p und

q konstant, ferner ist die Summe über die Produkte $S \cdot db$ gleich K, also wird

$$2N = \int K(2 - \cos p - \cos q)\, dg.$$

Hierin sind die gleichzeitigen Änderungen von g, p, q, vom Vorzeichen abgesehen, einander gleich, ferner hat g, wenn h die oben angegebene Bedeutung besitzt, von einem gewissen Anfangswerte g' bis $g' + h$ zu gehen, während sich p und q dabei in einem Intervall mit den Grenzen o und h bewegen. Daraus folgt

$$2N = K(2h - \sin h - \sin h),$$
$$N = K(h - \sin h),$$

womit sich für die Schnittdichte an der Stelle K der Ausdruck

$$h - \sin h$$

ergibt.

Ist dF das Flächenelement einer außerhalb A gezeichneten Figur B, ferner, wie vorhin, h der Winkel zwischen den von dF aus an A gezogenen Tangenten, so fällt von der Menge M_a in die Figur B die Schnittmenge

$$\int (h - \sin h)\, dF.$$

Dehnt man B über das ganze außerhalb A gelegene Gebiet aus, so erhält man die Schnittmenge M_a, also ist

$$U^2 - 2\pi J = 2M_a = 2\int (h - \sin h)\, dF,$$

wo die Integration über die ganze Fläche außerhalb A zu erstrecken ist. Die vorstehende elegante Formel rührt von *Crofton* her.

Es bedarf keiner besonderen Darlegung, wie sich die gefundenen Beziehungen in Sätze über W-Größen umdeuten lassen. Ferner sieht man, wie sich bei weiterer Verfolgung des eingeschlagenen Weges Aufgaben stellen lassen, bei denen Gerade und Ebenen im Raume mit krummen Oberflächen und Raumkurven in Verbindung gebracht werden. Da jedoch die gegebenen Beispiele genügen, um das Wesen der Methode deutlich zu machen, so verweise ich bezüglich weiterer Ausführungen auf das oben genannte Werk von *Czuber*.

Neunte Vorlesung.

Anwendungen und Fragestellungen.

§ 63. Mustert man die Aufgaben, die seither in den Lehrbüchern der W.-R. gewissermaßen den eisernen Bestand von Beispielen gebildet haben, so begegnet man neben den bisher behandelten Gegen-

ständen in der Hauptsache noch den Untersuchungen über Fehlertheorie, Versicherungswesen und Statistik. sowie über die Wahrscheinlichkeit von Zeugenaussagen und Abstimmungen. Nun hat die Behandlung der drei zuerst genannten Gegenstände im Laufe der Zeit einen Gang genommen, der sie weit über die Rolle einer bloßen Nutzanwendung der W.-R. hinaushebt; in diesen, zu selbständigen Disziplinen erwachsenen Untersuchungsgebieten ist die W.-R. nur ein Hilfsmittel, und zwar nur eines neben verschiedenen anderen. Die Sache liegt ähnlich, wie zwischen der Trigonometrie einerseits und der Geodäsie oder der sphärischen Astronomie andererseits: es wird niemand mehr einfallen, den Inhalt der zuletzt erwähnten Wissenschaften einem Lehrbuche der Trigonometrie als bloße Nutzanwendung anzuhängen. Aus diesem Grunde soll hier die Fehlertheorie und das Versicherungswesen beiseite bleiben, und von der mathematischen Statistik nur das Grundproblem berührt werden, weil damit ein natürlicher Übergang zu der Aufgabe der Kollektivmaßlehre gegeben ist. Zuvor möge jedoch die Wahrscheinlichkeit von Abstimmungen und Zeugenaussagen besprochen werden, da dieser Gegenstand in gewisser Hinsicht recht lehrreich ist.

Die Anwendung der W.-R. auf die Ergebnisse von Abstimmungen und Zeugenaussagen ist zuerst von *Condorcet*[1]) unternommen und von *Laplace* in der dritten Ausgabe seiner „Théorie analytique ...“ weiter verfolgt worden. Später hat dann noch *Poisson* in einem besonderen Werke[2]) den Gegenstand ausführlicher behandelt. Um deutlich zu machen, worauf es hierbei ankommt, diene folgendes einfache Beispiel.

Aus einer größeren Liste von Geschworenen, die kurz die *große Jury* heißen möge, wird für eine bestimmte Verhandlung eine *kleine Jury* von zwölf Mitgliedern ausgelost, die das verlangte Verdikt tatsächlich zu fällen hat. Wir wollen uns nun denken, daß der vorgelegte Fall von der großen Jury beurteilt werde, und daß hierbei jedes Mitglied seine Stimme mit „schuldig“ oder „nichtschuldig“ unabhängig, d. h. ohne vorhergegangenen Meinungsaustausch mit den anderen Mitgliedern, abzugeben habe. Das Ergebnis dieser Abstimmung, bei der S die Anzahl der „schuldig“ und N die der „nichtschuldig“ sein soll, wollen wir als das „maßgebende“ Urteil bezeichnen. Dann wird das Urteil der kleinen Jury, das wir uns auf dieselbe Weise, mit s „schuldig“ und n „nichtschuldig“, zustande gekommen denken, einerseits von den Zahlen S und N, andererseits von dem Ausfall der Losziehung, d. h. vom Zufall abhängen, also unter Umständen ganz anders, als das maßgebende Urteil, ausfallen können.

1) *Condorcet*, Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris 1785.

2) *Poisson*, Recherches sur la probabilité des jugements en matière criminelle et en matière civile. Paris 1837.

Da dieses Hineinspielen des Zufalls unleugbar einen Mangel der ganzen Einrichtung bedeutet, so kann man sich die Aufgabe stellen, den nachteiligen Einfluß der Losziehung genauer zu untersuchen und durch passende Festsetzung der für die Verurteilung vorgeschriebenen Majorität möglichst herunter zu drücken. Man gelangt damit zu einer klaren Aufgabe der W.-R., die sich auch sofort auf das Urnenschema reduzieren läßt, indem man fragt, wie groß ist die W., daß aus einer Urne, die S schwarze und N weiße Kugeln enthält, s schwarze und n weiße Kugeln gezogen werden, wenn die gezogene Kugel nicht zurückgelegt wird.

Sollen nun die Beziehungen und Formeln, die man auf dem eingeschlagenen Wege erlangen kann, praktisch verwertet werden, so ist zuerst die Vorfrage zu erledigen, ob die bei der Herleitung gemachten Voraussetzungen in der Wirklichkeit auch zutreffen. Hierbei kommt im wesentlichen nur die oben angeführte Annahme in Betracht, daß die einzelnen Stimmen unabhängig voneinander abgegeben werden sollen. Diese Annahme ist nun aber in Wirklichkeit nicht erfüllt, denn das Verdikt erfolgt stets nach einem mehr oder minder lebhaften Meinungsaustausch zwischen den Mitgliedern der kleinen Jury. Infolgedessen muß man bei den nicht ganz zweifelsfreien Fällen, derentwegen ja die in Rede stehende Untersuchung angestellt wird, mit dem Umstande rechnen, daß bei einzelnen Mitgliedern der Jury das Votum recht wesentlich davon abhängt, mit welchen anderen Mitgliedern sie durch das Los zusammengeführt worden sind. Oder mit anderen Worten: unter den zwölf gezogenen Kugeln sind solche, deren Farbe im voraus gar nicht bestimmt ist, sondern von der Farbe der anderen gezogenen Kugeln in einer mathematisch schlechterdings nicht zu erfassenden Weise abhängt. Damit kommt in die Frage ein Element hinein, das die ganze Rechnung für eine ernsthaft zu nehmende Anwendung unbrauchbar macht und weiter nichts übrig läßt, als eine hübsche Rechenübung.[1]

§ 64. Die vorstehenden Bemerkungen legen die Frage nahe, wie denn ein solcher Mißgriff — anders kann man die Sache nicht nennen — überhaupt zustande kommen konnte. Die erste Veranlassung dazu ist wohl in der rationalistischen Auffassung zu suchen, die einen hervorstechenden Zug an den geistigen Strömungen des achtzehnten Jahrhunderts bildet; es galt als ausgemacht, daß eine streng vernunftgemäße Gestaltung aller Verhältnisse des menschlichen Lebens nicht nur geboten, sondern auch erreichbar sei. Für diese Anschauung galt dann weiter die mathematische Analysis als das gegebene und unbedingt wirksame Hilfsmittel der Untersuchung bei allen denjenigen

1) Ein näheres Eingehen auf die *Poissonsche* Arbeit findet sich in der zutreffenden Kritik, die *von Kries* (a. a. O. S. 253 flg.) an der herkömmlichen Behandlung des vorliegenden Gegenstandes geübt hat.

Dingen, bei denen auch Zahl und Maß in Frage kommt. Von hier aus war nur noch ein kleiner Schritt, eine unbewußt vollzogene Unterschiebung, erforderlich, um zu der Vorstellung zu gelangen, daß eine Aufgabe, wie die vorhin besprochene, einer einwandfreien mathematischen Lösung fähig sei. Bei den Glücksspielen nämlich oder, allgemeiner gesprochen, bei den Vorgängen, die mit Sicherheit auf das Urnenschema reduziert werden können, hatten — das wußte man — die Begriffe Zufall und Wahrscheinlichkeit einen guten Sinn und führten bei weiterer Entwickelung auf Sätze, die die Probe der Erfahrung aushielten. Daraufhin vollzog sich nun, und zwar recht früh, die Unterschiebung, daß die Sätze der W.-R. schlechthin überall da anwendbar seien, wo überhaupt von Zufall und Wahrscheinlichkeit die Rede ist. Die Prüfung, ob denn die für die Zufälle und Wahrscheinlichkeiten eines Glücksspiels charakteristischen Merkmale jedesmal vorhanden seien, wurde einfach unterlassen. Recht bezeichnend hierfür ist ein Beispiel für die W. von Zeugenaussagen, das wir nunmehr noch berühren wollen.

Aus einer Urne, deren Kugeln fortlaufend die Nummern 1, 2, 3, ... tragen, wird einmal gezogen; ein Augenzeuge sagt aus, daß die Nummer k gezogen worden sei; gesucht wird die W. für die Wahrheit dieser Aussage. Bei der Behandlung der Aufgabe werden zunächst vier Fälle unterschieden, die dadurch entstehen, daß man vier, paarweise entgegengesetzte, Möglichkeiten kombiniert, nämlich 1) der Zeuge irrt oder irrt nicht, 2) der Zeuge lügt oder lügt nicht. Die vier Fälle sind dann: 1) Irren und Lügen, 2) Irren und Nichtlügen, 3) Nichtirren und Lügen, 4) Nichtirren und Nichtlügen. Des weiteren werden für diese vier Möglichkeiten schlankweg bestimmte W-Größen angesetzt, ohne Rücksicht auf die Frage, ob das überhaupt einen Sinn habe. Daß man für das Irren und Nichtirren eine bestimmte W. ansetzen dürfe, kann man allenfalls noch zugeben, streng genommen allerdings nur danu, wenn es sich dabei um rein psychophysische Vorgänge handelt, und selbst bei diesen liegen nach Ausweis der Erfahrungen, die man in der Psychophysik mit der sogenannten Methode der richtigen und falschen Fälle macht, die Verhältnisse sehr viel verwickelter, als bei den Ziehungen aus einer Urne. Dagegen vermag ich nicht zu verstehen, wie man das Lügen oder Nichtlügen mit dem Ziehen einer weißen oder schwarzen Kugel in Parallele stellen kann — man müßte denn ein Individuum vor sich haben, das auf äußere Anreize hin rein reflektorisch mit Lügen oder Nichtlügen reagiert. Solche Individuen können allerdings vorkommen, man pflegt sie aber nicht als geistig normal zu betrachten.

Es kann nicht zweifelhaft sein, daß Beispiele, wie die genannten, aus den Darstellungen der W.-R. entweder einfach auszumerzen oder doch höchstens aus historischen Rücksichten als abschreckende Fälle

mitzunehmen sind. Selbstverständlich ist damit nicht gesagt, daß nun etwa auch statistische Aufzeichnungen, z. B. über die Verdikte der Schwurgerichte, wertlos seien, denn statistische Angaben sind beobachtete Tatsachen, die unter Umständen und bei richtiger Deutung wichtige Aufschlüsse geben können. Nur muß man sich davor hüten, jede Zahl, die bei einer statistischen Untersuchung als eine rH. auftritt, zugleich auch als eine 𝔚-Größe zu betrachten. Denn die Bildung der 𝔚-Größen setzt stets das Vorhandensein einer Reihe gleichmöglicher Fälle voraus, die einander ausschließen und voneinander unabhängig sind; es müssen also jedesmal ausreichende Gründe für die Zulassung einer solchen Voraussetzung vorliegen.

§ 65. Unter den Anwendungen der W.-R. waren oben auch Fragen aus dem Gebiete der Statistik genannt worden. Die hierher gehörenden Aufgaben besitzen das gemeinsame Merkmal, daß die zu untersuchenden 𝔚-Größen von unbekannten Ursachen abhängen, also aus den beobachteten Ergebnissen gewisser Versuche abgeleitet werden müssen. Damit gelangen wir zu denjenigen Anwendungen der W.-R., die wir früher als eine besondere zweite Klasse von den Fällen mit bekannter Ursache abgetrennt hatten. Die *Bayesschen* Sätze, die dabei herangezogen zu werden pflegen, haben wir bereits in der dritten Vorlesung (§ 19 flg.) kennen gelernt; ihre praktische Brauchbarkeit ist an die Voraussetzung geknüpft, daß man es mit „großen Zahlen“ oder, um es bezeichnender auszudrücken, mit *Massenbeobachtungen* zu tun habe. Die Wichtigkeit dieser Bedingung ist übrigens unmittelbar einleuchtend, denn wenn z. B. das Füllungsverhältnis einer Urne, die weiße und schwarze Kugeln enthält, aus Ziehungen mit zurückzulegender Kugel ermittelt werden soll, so ist niemand darüber im Zweifel, daß einige wenige Züge höchstens die Gewißheit von dem Vorhandensein der gezogenen Farben geben, dagegen über das Füllungsverhältnis nur eine ganz unsichere Aussage zulassen.

Massenbeobachtungen von der hier in Betracht kommenden Art wird man vorzugsweise dort zu suchen haben, wo umfangreiche Reihen von Einzelfällen nach „statistischer“, d. h. rein abzählender, Methode beobachtet und aufgezeichnet werden, wie das namentlich von der eigentlichen Statistik gilt, die die wechselnden Gestaltungen des menschlichen Lebens ziffernmäßig zu verfolgen unternimmt. Auch ist der Gedanke, auf solche Massenbeobachtungen die Prinzipien der W.-R anzuwenden, schon sehr früh aufgetaucht. Das bekannteste Beispiel aus der älteren Literatur ist die von *Halley* für Zwecke der Rentenversicherung entworfene Absterbeordnung[1]), die sich auf die viel ge-

1) *Halley*, An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw, with an attempt to ascertain the price of annuities upon lives. Philosophical Transactions, XVII, 1693.

nannten Breslauer Tafeln des Theologen *Kaspar Neumann* stützte. Diese Tafeln enthielten die monatweise geordneten, nach Alter und Geschlecht gesonderten Zahlen für die Bewegung der Breslauer Bevölkerung während der Jahre 1687—1691. Selbstverständlich ist nicht alles, was hierbei statistisch beobachtet und berechnet wird, dazu angetan, Gegenstand der W.-R. zu werden: es wird z. B. niemandem ernstlich einfallen können, aus dem Menschenverlust, den die Kriege vergangener Jahre gebracht haben, wahrscheinliche Werte für zukünftige Verluste dieser Art berechnen zu wollen. Eine lehrreiche Erörterung über die möglichen Beziehungen zwischen der W.-R. und der Statistik findet sich in einer Abhandlung von W. *Lexis*[1]), aus der nachstehend das für unseren Zweck Wesentliche angeführt werden soll. Obgleich die Untersuchung von *Lexis* zunächst nur die Vorgänge in der menschlichen Gesellschaft zum Gegenstand nimmt, so gilt sie — mutatis mutandis — doch auch über dieses Gebiet hinaus.

§ 66. Der Zustand der menschlichen Gemeinschaft — so beginnt *Lexis* — wird einerseits bedingt durch die positiven Gestaltungen und Normen der Gesellschaft und des Staates, die historisch geworden sind, und deren Änderungen historische Ereignisse bilden; andererseits aber auch durch das gewöhnliche, relativ stetige Tun und Leiden der Individuen in ihrer mannigfaltigen Gruppierung, das in seinen einzelnen Elementen nicht festgehalten werden kann, aber charakteristische, der wissenschaftlichen Beobachtung zugängliche *Massenerscheinungen* erzeugt. Die Statistik hat die selbständige Aufgabe, diese Massenerscheinungen des Menschenlebens nach exakter Methode aufzufassen und zu untersuchen. Sehr verfehlt wäre es jedoch, wenn man *alle* menschlichen Massenerscheinungen lediglich vom statistischen Gesichtspunkte betrachten wollte. Denn viel wichtiger als die Aufhebung des Einzelereignisses in einer konkreten statistischen *Summe* ist die Aufhebung desselben in einer begrifflichen *Verallgemeinerung*. Wenn die Einzelfälle nur Exemplare einer und derselben *Gattung* sind, und wenn wir ferner diese Gattung des Geschehens aus einer Ursache oder einem Ursachensystem begreifen können, so ist dieser abstrakte *Begriff* des Ereignisses wissenschaftlich von größerem Interesse, als die Zählung seines konkreten Vorkommens. Eine solche begriffliche oder *generische* Auffassung der menschlichen Massenerscheinungen aber ist namentlich dann möglich, wenn wir, gestützt auf psychologische Erwägungen, Selbstbeobachtungen oder alltägliche Erfahrungen, in jedem Einzelereignis die gleiche überwiegend wirksame Ursache, insbesondere also in jeder zu der Masse beitragenden Einzelhandlung die gleiche durchschlagende Triebfeder zu erkennen vermögen.

1) *W. Lexis,* Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft. Freiburg i. B., Fr. Wagner, 1877.

Es ist nicht nötig, Beispiele hierzu anzuführen, denn die Lehre von den wirtschaftlichen und sozialen Dingen stellt sich von Anfang an die Aufgabe, die kausalen Beziehungen aufzudecken, die zwischen den zu untersuchenden Vorgängen bestehen. Die statistischen Zahlenreihen sind dabei nur Mittel zum Zweck: sie dienen dazu, den geschichtlichen Verlauf von Massenerscheinungen deutlich zu beschreiben, sie dienen ferner, ähnlich wie der Vorlesungsversuch des Physikers, dazu, erkannte Beziehungen zu demonstrieren, endlich können sie auch unter der Form statistischer „Reaktionen“ benutzt werden, um vermutete Beziehungen als erkennbar oder nichterkennbar nachzuweisen. In allen diesen Fällen wird die statistische Ziffer nur dadurch wissenschaftlich fruchtbar, daß wir sie mit unseren sonstigen Erfahrungen über die Natur des gesellschaftlichen und wirtschaftenden Menschen verbinden.

Bei den *generischen* Massenerscheinungen liegt also das Wesentliche darin, daß die Einzelfälle als Exemplare einer bestimmten Gattung von Vorgängen aufgefaßt und als solche aus gewissen, unserer Einsicht zugänglichen Ursachen erklärt werden, daß dagegen die statistische Ermittelung der hierbei auftretenden numerischen Verhältnisse lediglich als Hilfsmittel der Untersuchung dient. Über dieses Gebiet hinaus gibt es nun aber noch Massenerscheinungen, deren wissenschaftliches Interesse zunächst nur in ihren numerischen Verhältnissen liegt. Wenn man z. B. sagt: „von den Geborenen einer gewissen Zeitstrecke sterben viele vor Ablauf des ersten Lebensjahres“, so ist mit diesem Satze nicht viel anzufangen. Wesentlich anders stellt sich dagegen die Sache, wenn es heißt: „der Prozentsatz der Gestorbenen ist soundsogroß und unter normalen Verhältnissen für eine Reihe von aufeinander folgenden Generationen merklich konstant“. Denn wir erhalten dann einen vergleichsweise allgemeinen Erfahrungssatz, dessen Wichtigkeit auch nicht durch den Umstand beeinträchtigt wird, daß wir die beobachteten Zahlenwerte als gegebene, aus anderen bekannten Tatsachen nicht herzuleitende Größen hinzunehmen haben.

Lexis bezeichnet nun die zuletzt genannten Massenerscheinungen als *konkrete* und betrachtet ihre Gesamtheit als das Gebiet, auf dem die Statistik als Wissenschaft von den menschlichen Massenerscheinungen selbständig aufzutreten vermag. Als Hilfsmittel der Untersuchung dient dabei die W.-R., deren Sätze bei den generischen Massenerscheinungen höchstens ausnahmsweise Anwendung finden können, weil ja da, wo die Einsicht in die Ursachen eines Vorganges den Anschein der Zufälligkeit aufhebt, auch die Anwendung der Zufallstheorie gegenstandslos wird.

§ 67. Die selbständigen Ergebnisse der Statistik bestehen vorzugsweise darin, daß sie die genäherte Konstanz gewisser numerischer

Verhältnisse der Massenerscheinungen feststellt. Diese Konstanz drängte sich schon früh der Wahrnehmung auf und heischte eine Erklärung, wenn man nicht etwa dem Vorbilde von *Süßmilch* folgen wollte, der in den beobachteten numerischen Regelmäßigkeiten den Ausdruck einer *göttlichen Ordnung* sah und damit das Suchen nach einem befriedigenden Erklärungsgrunde kurzweg abschnitt. Sieht man sich nun die Massenerscheinungen, die ausgesprochene numerische Regelmäßigkeiten zeigen, genauer daraufhin an, ob bei ihnen offenliegende Gründe für die gedachte Eigenschaft zu erkennen sind oder nicht, so gelangt man zu der Einsicht, daß zwei Gruppen zu unterscheiden sind. Einmal nämlich gibt es Massenerscheinungen, deren Einzelfälle in einem bestimmten Zusammenhange miteinander stehen, und zwar derart, daß die durch gewisse Einzelfälle erzeugte Abweichung von der numerischen Regelmäßigkeit durch eine entgegengesetzte Abweichung anderer Einzelfälle ständig ausgeglichen wird. Wenn z. B. in einem gewissen Lande ein umfangreicher Industriezweig besteht, der zweier aus verschiedenen Ländern einzuführender Rohstoffe A und B bedarf und für das fertige Produkt von dem Stoffe A immer doppelt soviel als von B verarbeiten muß, so ist es niemandem auffällig, daß ein relativer Überschuß in der Einfuhr von A oder B nach einiger Zeit seine Ausgleichung durch ein entsprechendes relatives Zurückbleiben der Einfuhr findet, daß sich also im Durchschnitt aus einem längeren Zeitraum die Einfuhrmengen von A und B ständig sehr nahe in dem Verhältnis von 2 zu 1 stellen. *Lexis* bezeichnet solche Massenerscheinungen als *verbundene* im Gegensatz zu den *unverbundenen*, deren Einzelfälle keinen, auf die Herbeiführung eines bestimmten numerischen Verhältnisses hinzielenden, Zusammenhang erkennen lassen.

Hiernach bedarf die Frage, wie man sich das Entstehen der numerischen Regelmäßigkeiten zu denken habe, einer weiteren Erörterung nur noch für den Fall der unverbundenen Massenerscheinungen. Man könnte sich nun zunächst denken, daß dabei eine naturgesetzliche Beziehung obwalte, die jene Regelmäßigkeiten mit Notwendigkeit herbeiführt, etwa in derselben Weise, wie dem fallenden Stein durch eine bestimmte Ursache am Ende der ersten Sekunde eine Geschwindigkeit von rund 10 m erteilt wird. Eine solche Auffassung hat in der Tat Vertreter gefunden; bezeichnend in dieser Hinsicht ist die vielzitierte Phrase *Quetelets* von dem „Budget des Schaffots und der Gefängnisse“, die, folgerecht weiter gedacht, zu dem Schlusse führt, daß in der letzten Woche eines Jahres soundsoviele Individuen sich noch schleunigst das Leben nehmen müssen, falls etwa die vorhergehenden 51 Wochen des Jahres noch nicht die vom Statistiker herausgerechnete Normalzahl der Selbstmorde eingebracht haben sollten. Es ist nicht nötig, auf diese Vorstellung von einem naturgesetzlichen

Ursprung der statistischen Regelmäßigkeiten näher einzugehen, denn sie hat sich bisher als völlig unfruchtbar erwiesen, und man ist mit ihr in Wahrheit nicht über die nackte Aussage von dem Vorhandensein solcher Regelmäßigkeiten hinausgekommen.

Eine andere Auffassung geht davon aus, daß man es bei der Konstanz statistischer Ergebnisse mit derselben Erscheinung zu tun habe, die sich bei der hinreichend häufigen Wiederholung von Glücksspielen einstellt und kurz als Ausgleichung des Zufalls bezeichnet wird; wenn man z. B. einen fehlerfreien Würfel tausendmal wirft, so ergibt sich die rH. der Einerwürfe sehr nahe gleich dem Bruche 1 : 6, und diese Annäherung wiederholt sich, wenn die Versuchsreihe wiederholt wird. Im Grunde genommen ist die Vorstellung, daß man die unverbundenen Massenerscheinungen ohne weiteres auf das Urnenschema reduzieren dürfe, ebenso alt, wie die ersten Anwendungen der W.-R. auf die Statistik, d. h. nur wenig jünger, als die erste Begründung der Zufallstheorie; man betrachtete dabei die Einzelfälle einer Massenerscheinung als Dinge, die in der gleichen Weise wie das Rollen eines Würfels, dem Zufall unterworfen seien, und sah demgemäß in der Zurückführung auf das Urnenschema eine selbstverständliche, für die Rechnung als Ausgangspunkt zu nehmende Voraussetzung. *Lexis* setzt nun — und das ist der eigentliche Kern seiner Ausführungen — an die Stelle dieser früher ohne besondere Begründung eingeführten *Annahme* die *Aufgabe*, jedesmal zu untersuchen, ob und wieweit denn die Heranziehung des Urnenschemas überhaupt zulässig sei.

Die möglichen Ergebnisse einer solchen Untersuchung können, falls das Beobachtungsmaterial überhaupt sichere Schlüsse zuläßt, die folgenden sein. Zunächst kann der Fall eintreten, daß die vorgelegten Zahlen sich wie die Ergebnisse eines Massenspiels verhalten, dessen Chancen entweder konstant sind oder doch nur nach Zufall um feste Mittelwerte herumschwanken. Als Beispiele von derartigen *typischen* Reihen behandelt *Lexis* das Geschlechtsverhältnis der Geburten und die Frage nach dem „Normalalter" einer Generation. Weiter kann der Fall eintreten, daß die vorgelegte Reihe dem Schema eines Glücksspiels folgt, dessen Chancen mit nicht-zufälligen Änderungen fortschreitender oder oszillierender Art behaftet sind. Solche *symptomatische* Reihen bilden aller Wahrscheinlichkeit nach die Hauptmenge der unverbundenen menschlichen Massenerscheinungen. Schließlich kommt noch der Fall in Betracht, daß sich die beobachteten numerischen Regelmäßigkeiten nachweisbar anders als die Ergebnisse eines Massenspiels verhalten: man wird dann die Frage zu untersuchen haben, ob nicht etwa die vorgelegte Massenerscheinung eine versteckt liegende Verbundenheit besitze.

Die Durchführung der verlangten Untersuchung setzt voraus, daß man über Kriterien verfügt, vermittels deren man erkennen kann,

welcher von den vorhin genannten Fällen des Urnenschemas bei einer gegebenen statistischen Reihe jedesmal vorliegt. Ich will hier auf diese Kriterien für den Augenblick nicht näher eingehen, weil der Gegenstand weiterhin im Zusammenhange mit anderen Dingen von einem allgemeineren Gesichtspunkte aus zu behandeln sein wird, und wende mich jetzt, indem ich die *Lexis*schen Darlegungen verlasse, zu einer Fragestellung, die von *G. Th. Fechner* in einem nachgelassenen Werke[1]) eingehend untersucht worden ist. Der Deutlichkeit halber beginne ich dabei mit einem einfachen Beispiel.

§ 68. Man denke sich für einen gewissen Zeitraum von den Rekruten eines bestimmten Aushebungsbezirks und einer bestimmten Altersklasse die Körperlängen x gegeben, wobei die Werte der x auf volle Zentimeter abgerundet sein mögen, so daß z. B. der Wert $x = 171$ cm allen denjenigen Individuen zukommt, deren genaue Körperlänge zwischen 170.5 und 171.5 cm liegt. Die Gesamtheit der gemessenen Individuen mit ihren x bezeichnen wir nach *Fechner* als einen *Kollektiv-Gegenstand* (kurz K.-G.) oder auch als eine *Kollektiv-Reihe*, mit dem Vorbehalt, die vollständige Bedeutung dieses Wortes nachher noch genauer abzugrenzen. Die einzelnen Individuen bilden dann die *Glieder* der Kollektivreihe, während die Gliederanzahl m den Namen *Umfang* des K.-G. erhalten soll.

Werden die x der einzelnen Glieder in der Reihenfolge hingeschrieben, in der sie gemessen worden sind, so erhält man die *Urliste*. In dieser Liste werden, falls die Glieder des K.-G. nicht etwa schon vor der Messung in bestimmter Weise geordnet wurden, die einzelnen Zahlen keinerlei Gang oder Gesetz erkennen lassen, sondern im allgemeinen regellos hin- und herspringen. Anders stellt sich — ein genügend großes m vorausgesetzt — dagegen die Sache, wenn man die x nach ihrer Größe ordnet oder nach *Fechners* Ausdruck aus der Urliste die *primäre Verteilungstafel* herleitet. Zu dem Ende schreiben wir die x nach der Reihenfolge ihrer Größe in einer Spalte untereinander, setzen jedoch ein mehrfach vorkommendes x immer nur einmal an. Ferner schreiben wir neben jedes x in einer zweiten Spalte die Zahl y hin, die angibt, wie oft das betreffende x in der Urliste vorkommt. Die Summe der y ist dann offenbar gleich dem Umfange m.

Die so hergeleitete Verteilungstafel wird nun bei einem genügend großen m folgendes Verhalten zeigen. Die y sind anfangs klein, steigen dann mit gelegentlichen kleinen Sprüngen zu einem Maximum an, um nachher in ähnlicher Weise wieder abzufallen. Noch deutlicher wird dies, wenn man die Beziehung zwischen den x und y geometrisch

1) *G. Th. Fechner*, Kollektivmaßlehre. Herausgegeben von *G. F. Lipps*. Leipzig, W. Engelmann, 1897.

darstellt. Zu dem Ende bemerken wir zunächst, daß man, ohne den Sinn und Inhalt der Verteilungstafel zu ändern, in der x-Spalte beliebig viele, in der vorgelegten Kollektivreihe nicht vorkommende x hinzufügen darf, sobald man nur das entsprechende y gleich Null setzt. Bezeichnet man ferner die x als *leer* oder *voll*, je nachdem das zugehörige y null ist oder nicht, so darf man nach dem Gesagten voraussetzen, daß nach Hinzufügung der nötigen leeren x die Zahlen der x-Spalte von $-\infty$ bis $+\infty$ laufen. Dies vorausgeschickt tragen wir die nach ganzen Zentimetern fortschreitenden x als Abszissen, die entsprechenden y als Ordinaten ab, verbinden die Endpunkte der aufeinander folgenden Ordinaten durch gerade Linien, und erhalten so einen gebrochenen Linienzug, den wir kurz als die Kurve des vorgelegten K.-G. bezeichnen wollen. Diese Kurve verläuft anfangs in der Abszissenachse, steigt dann, von Sprüngen abgesehen, zu einem Maximum auf, um darauf in derselben Weise wieder abzufallen und schließlich in der Abszissenachse weiter zu laufen. Betrachtet man nun den Linienzug nach seinem Verhalten im ganzen, so gelangt man zu dem Satze, daß sein Verlauf, trotz der erwähnten und von unausgeglichenen Zufälligkeiten herrührenden Sprünge, deutlich Regel und Gesetz erkennen läßt, und zwar in demselben Sinne, in dem man z. B. bei der Erdgestalt, trotz der handgreiflichen Gegensätze zwischen Berg und Tal, von einer Kugel oder einem Ellipsoid spricht.

Die soeben als Beispiel benutzte Rekrutentafel mit ihrer Kurve ist ein willkürlich herausgegriffener Fall unter zahllosen anderen, die auf dem Gebiete der Massenerscheinungen anzutreffen sind, und die in der gleichen Weise wie jene Tafel auf regelmäßig verlaufende Kurven führen. Der charakteristische Berg mit zweiseitigem Abfall findet sich, wie *Fechner* u. a. an den Kurven für die Abmessungen von Galleriegemälden nachweist, sogar bei solchen Dingen, die man lediglich als Produkte der menschlichen Willkür anzusehen gewohnt ist.

Die Tatsache, daß regelmäßig verlaufende Verteilungskurven bei den verschiedensten Dingen nachgewiesen werden können, ist nun für die W.-R. von grundlegender Bedeutung, denn sie liefert das Fundament für den Satz, daß in der Wirklichkeit näherungsweise das stattfindet, was wir früher als Ausgleichung des Zufalls oder — genauer gesprochen — als die gleichmäßige Erschöpfung der gleichmöglichen Fälle bezeichnet haben. Ohne diesen Satz würde man angesichts der genannten Tatsache zu der Vorstellung genötigt sein, daß zwischen den Gliedern jeder Kollektivreihe eine schlechthin unverständliche Verbundenheit bestehe, die auf die Erzeugung der beobachteten Regelmäßigkeiten hinarbeitet.

Des weitern ist der Umstand hervorzuheben, daß die Beispiele für die Ausgleichung des Zufalls nicht erst mühsam an versteckten Stellen aufgesucht oder durch besondere Experimente herbeigeschafft werden müssen, sondern sich ungesucht auf Schritt und Tritt der

Wahrnehmung aufdrängen. Es wäre um die Anwendung der W.-R. nicht sonderlich gut bestellt, wenn die Berechtigung dazu auf einige ad hoc angestellte Versuchsreihen mit Würfeln und dergleichen gestützt werden müßte; solche Versuche sind zur Erläuterung allgemeiner Sätze sehr nützlich, dagegen ist die Berechtigung zur Formulierung solcher Sätze auf eine breitere Erfahrungsgrundlage zu stellen, als sie einzelne Versuche gewähren können.

§ 69. *Fechner* knüpft an die genannte Erfahrungstatsache die grundlegende Frage, ob es nicht möglich sei, bei den in Rede stehenden Kurven gemeinsame stets wiederkehrende Maßbeziehungen aufzudecken. Ist diese Frage zu bejahen, so erhält man damit offenbar die Bausteine zu einer allgemeinen *Formenlehre* der Kollektivgegenstände. Diese Fragestellung, sowie der damit verbundene ernsthafte Versuch, eine Formenlehre der Kollektivreihen wirklich aufzubauen, bilden ein wesentliches Verdienst des *Fechner*schen Buches. Dieses Verdienst wird auch nicht dadurch geschmälert, daß einzelne besondere Kollektivreihen, z. B. Reihen von Beobachtungsfehlern schon lange vor *Fechner* mehr oder minder eingehend untersucht worden sind, und daß der von *Fechner* benutzte mathematische Apparat ein höchst primitiver ist. Denn der mathematische Apparat hat sich nachher ohne besondere Schwierigkeiten verbessern lassen, während die als Hauptsache anzusehende Forderung, das ganze Gebiet von einheitlichen und umfassenden Gesichtspunkten aus zu bearbeiten, meines Wissens zuerst von *Fechner* mit voller Bestimmtheit erhoben worden ist.

Bei *Fechner* erscheint die Kollektivmaßlehre als ein selbständiges Kapitel der angewandten Mathematik *neben* der Wahrscheinlichkeitsrechnung. Diese Scheidung ist indessen nicht notwendig. Faßt man nämlich die W.-R. als eine „Häufigkeitsrechnung“ auf, so ist das Gebiet ihrer Anwendungen durch die Massenerscheinungen gegeben. Das Objekt der Untersuchung ist also dasselbe, wie in der Kollektivmaßlehre. Dementsprechend wird die weiterhin folgende Darstellung lehren, *daß man es in der Kollektivmaßlehre lediglich mit dem folgerechten Ausbau der Methoden der W.-R. zu tun hat.*[1]) So wird sich z. B. zeigen, daß die Sätze von *Bernoulli* und *Poisson*, die in der herkömmlichen Darstellung ziemlich unvermittelt auftreten, lediglich einfache Grenzfälle von theoretisch konstruierten Kollektivreihen sind und sich als solche ungezwungen in den allgemeinen Gedankengang einfügen.

Bei der nunmehr folgenden Behandlung der Kollektivmaßlehre soll nachstehende Gliederung des Stoffes innegehalten werden. Zunächst

1) Daß diese Auffassung auch von anderen Seiten geteilt wird, lehrt u. a. der Aufsatz von *G. Helm* „Die Wahrscheinlichkeitslehre als Theorie der Kollektivbegriffe“ (Annalen der Naturphilosophie, Band I).

ist der Begriff eines K.-G. schärfer, als oben vorläufig geschehen war, abzugrenzen. Daran schließt sich die Entwicklung der mathematischen Hilfsmittel, die für die ziffermäßige Untersuchung der K.-G. notwendig sind. Der Gebrauch dieser Hilfsmittel wird dann zuerst an theoretisch konstruierten Kollektivreihen erläutert, während die Erörterung über die zweckmäßige Behandlung beobachteter Kollektivgegenstände den Schluß der Darstellung bildet.

Zehnte Vorlesung.

Grundbegriffe der Kollektivmaßlehre.

§ 70. Der erste Schritt, den wir bei der Darstellung der Kollektivmaßlehre auszuführen haben, besteht in der schärferen Begriffsbestimmung der Kollektivgegenstände. Wir wollen dabei an den oben benutzten Fall einer Rekrutentafel anknüpfen, und zwar in der Weise, daß wir die verschiedenen Merkmale des Beispiels ihrer konkreten Besonderheit entkleiden. Wir dürfen dann sagen, daß es sich bei einem K.-G. zunächst um eine gewisse Vielheit von Dingen handele, die in bestimmten Merkmalen übereinstimmen und deshalb als gleichartig angesehen werden können. So sind z. B. in der Rekrutentafel die die Gleichartigkeit begründenden oder *konstanten* Merkmale durch die Begriffe „normal entwickelter Mensch“, „Geschlecht“, „Altersstufe“ und „Aushebungsbezirk“ gegeben. Außer den konstanten Merkmalen sind nun aber noch andere vorhanden, die wir die *veränderlichen* nennen, weil sie bei dem vorgelegten K.-G. von einem Gliede zum andern wechseln können. Von diesen veränderlichen Merkmalen wird eines (die Körperlänge) herausgegriffen und zum Ordnen der Urliste benutzt, wodurch sich die vorgelegte Urliste in eine primäre Verteilungstafel umwandelt. Die vorgenommene Ordnung kann als eine *statistische* bezeichnet werden, weil dabei die Glieder, die in dem ausgewählten veränderlichen Merkmal übereinstimmen, nicht weiter unterschieden, sondern lediglich nach der Häufigkeit ihres Vorkommens gezählt werden. An der Hand dieser Bemerkungen wird es jetzt unmittelbar verständlich sein, wenn wir sagen: *Ein Kollektivgegenstand ist eine Vielheit von gleichartigen Dingen, die nach einem veränderlichen Merkmal statistisch geordnet werden kann.*

Das ordnende Merkmal wollen wir, sobald eine numerische Maßbestimmung seiner Abstufungen vorliegt, auch als das *Argument* des K.-G. bezeichnen. Da an konkreten Dingen im allgemeinen unendlich viele Merkmale nachgewiesen werden können, so lassen sich bei einem vorgelegten K.-G. statt eines einzigen ordnenden Arguments auch

mehrere gleichzeitig ins Auge fassen. Handelt es sich z. B. um die Untersuchung einer Vielheit von Eheschließungen, so wird die Altersstufe der Frau dabei ebenso wichtig sein, wie die des Mannes; die endgültige Ordnung dieser Vielheit nach Alterstufen wird also zwei Argumente gleichzeitig zu berücksichtigen haben. Es ist indessen nicht nötig, bei dem Fall mehrerer Argumente zu verweilen, denn seine Behandlung läßt sich auf den Fall nur eines Argumentes zurückführen.

Die Worte „Vielheit" und „gleichartig" in der aufgestellten Definition der K.-G. enthalten eine gewisse Unbestimmtheit, die jedoch in der Natur der Sache liegt. Da die Untersuchung eines K.-G. darauf ausgeht, das Vorhandensein oder Fehlen einer Regelmäßigkeit nachzuweisen, so muß zunächst der Umfang des K.-G. groß genug sein, um einen vertrauenswürdigen Schluß nach der einen oder anderen Richtung hin zu gestatten, weil sonst die Untersuchung von vornherein zur Ergebnislosigkeit verurteilt ist. Wie groß nun aber der Umfang des K.-G. für den gedachten Zweck mindestens sein muß, das hängt ganz wesentlich von den Umständen ab; es kann vorkommen, daß schon ein halbes Hundert von Gliedern ein brauchbares Resultat liefert, während in anderen Fällen Tausende nötig sind.

Die vorgeschriebene „Gleichartigkeit" der Glieder fordert, daß letztere durch einen Art- oder Gattungsbegriff zusammengehalten werden; es läßt sich aber nicht allgemein im voraus festsetzen, wie weit oder wie eng der Kreis der konstanten Merkmale zu ziehen sei, um mit einiger Sicherheit auf ein brauchbares Ergebnis der Untersuchung rechnen zu können. Daß man z. B. bei einer Tafel der Körperlängen nicht Kinder und Erwachsene zusammenwerfen dürfe, erscheint ja unmittelbar einleuchtend, dagegen läßt sich ohne nähere Untersuchung nicht angeben, wie weit oder wie eng man bei einer solchen Tafel die Grenzen einer Altersstufe anzusetzen habe. Wird bei dem Sammeln einer Kollektivreihe eine möglichst weitgehende Gleichartigkeit, d. h. eine möglichst große Anzahl von konstanten Merkmalen vorgeschrieben, so gewinnt ein derart gebildeter K.-G. natürlich an Bestimmtheit, dafür wächst aber auch die Schwierigkeit, ein umfangreicheres Material zusammenzubringen. Man hat es also bei einem beobachteten K.-G. für gewöhnlich mit einem Kompromiß zwischen den beiden einander widerstreitenden Forderungen möglichster Gleichartigkeit und möglichst großen Umfanges zu tun.

§ 71. Zu der von uns aufgestellten Definition ist noch zu bemerken, daß *Fechner* bei seinen Untersuchungen eine merklich engere Begriffsbestimmung zugrunde legt. Er setzt zunächst voraus, daß das veränderliche Merkmal *unmittelbar* nach Zahl und Maß bestimmt sei, und daß die Gesamtheit der möglichen Argumentwerte eine stetige Mannigfaltigkeit bilde. Diese Einschränkungen sind entbehrlich.

Beispielsweise könnte man auf Grund einer passenden Farbenskala zu der Farbe der Haare oder Augen bei den Schulkindern einer Großstadt sehr wohl eine Verteilungstafel konstruieren, indem man die einzelnen Farbentöne der Skala durch Nummern unterscheidet und mit letzteren rechnet. Ebenso hindert uns nichts, die einzelnen Spalten in der zehnstelligen Logarithmentafel von *Vega* als die Glieder eines K.-G. zu behandeln und als Argument z. B. die nur auf ganze Zahlen führende Häufigkeit der Endnullen zu wählen. Es ist sehr merkwürdig, daß *Fechner* mit keiner Silbe erwähnt, weshalb er die K.-G. mit diskreten Argumenten beiseite läßt; die von ihm gewählte Form der mathematischen Behandlung bietet zwar bei den diskreten Argumenten gewisse Schwierigkeiten, aber keine unübersteiglichen Hindernisse, ferner lassen sich gerade für diesen Fall gute Beispiele mit vergleichsweise geringer Mühe herbeischaffen.

Eine andere von *Fechner* gemachte Einschränkung besteht in der Voraussetzung, daß in der Urliste das Argument von einem Gliede zum anderen nach Zufall variiere, daß also, wie es *Fechner* nennt, in der Urliste keine *Sukzessionsabhängigkeit* vorhanden sei. Auch diese Einschränkung ist entbehrlich. Ob zwischen den aufeinander folgenden Gliedern der Urliste ein Zusammenhang besteht oder nicht, das ist, wie wir sehen werden, für die mathematische Behandlung zunächst gleichgültig und wird erst dann von Belang, wenn es sich um die *Deutung* der errechneten Zahlengrößen handelt.

Schließlich mag noch hervorgehoben werden, daß die von uns gewählte Definition auch solche K.-G. umfaßt, die nicht beobachtet, sondern willkürlich ersonnen sind. Der Nutzen, den man aus diesem Umstande ziehen kann, wird weiterhin ersichtlich werden.

§ 72. Nachdem wir den Begriff des K.-G. in der für unsere Zwecke passenden Weise festgesetzt haben, ist jetzt die Einrichtung der Verteilungstafel und ihre geometrische Darstellung zu erörtern, wobei als Vorbereitung einige Bemerkungen über das Argument vorauszuschicken sind.

Die Abstufungen des „ordnenden“ veränderlichen Merkmals denken wir uns stets ziffermäßig ausgedrückt. Diese Ausdrucksart ergibt sich von selbst, wenn das veränderliche Merkmal eine unmittelbar nach Zahl und Maß bestimmte Eigenschaft, z. B. eine Länge, ein Gewicht, eine Temperatur, eine relative Häufigkeit oder dergleichen bezeichnet. Aber auch dann, wenn die Abstufungen des veränderderlichen Merkmals zunächst als rein qualitative Unterschiede auftreten, wie z. B. bei der oben erwähnten Untersuchung der Haarfarbe, ist es immer zulässig, die einzelnen Stufen durch Nummern zu unterscheiden, also ziffermäßig auszudrücken. Zu jedem Gliede eines K.-G. gehört dann also stets ein bestimmter Zahlenwert x des Arguments. Die Werte der x denken wir uns immer zugleich auch als Abszissen

auf einer Abszissenachse abgetragen und werden demgemäß auch kurz von den Punkten x sprechen.

Die Argumentwerte x, die in der Urliste eines vorgelegten K.-G. entweder tatsächlich vorkommen oder aber den Umständen nach vorkommen könnten, wollen wir als die *möglichen* Werte bezeichnen, um sie von den *außermöglichen* Werten zu unterscheiden, die die Größe x, als reelle unbeschränkte Veränderliche gedacht, auch sonst noch zwischen den Grenzen $\pm \infty$ annehmen kann. Man denke sich z. B. aus einer längeren Reihe von Geburten einen K.-G. gebildet, indem man Gruppen von je hundert oder je tausend Geburten abteilt, diese als die Glieder des K.-G. ansieht und als veränderliches Merkmal das Auftreten der Knabengeburten auswählt. Nimmt man nun als Argument die relative Häufigkeit x der Knabengeburten, so sind die möglichen x auf das Intervall von Null bis Eins beschränkt. Wählt man dagegen, wie das in der Statistik öfters geschieht, als Argument das Geschlechtsverhältnis oder den Quotienten $x : (1 - x)$, so dehnt sich das Intervall der möglichen Argumente über das ganze Gebiet der positiven Zahlen aus.

Wie das vorstehende Beispiel lehrt, besteht für die Wahl des Arguments, d. h. für die Maßbestimmung des veränderlichen Merkmals, ein gewisser Spielraum. Man darf für das ursprünglich gewählte Argument x irgend ein neues Argument x' einführen, sobald nur die Beziehung zwischen x und x' so beschaffen ist, daß zu verschiedenen x stets verschiedene x' gehören und umgekehrt. Dehnt sich also z. B. das Gebiet der möglichen x bis zu den Grenzen $\pm \infty$ aus, so würde, wenn als neues Argument x' die Funktion $\operatorname{arctg} x$ eingeführt wird, das Gebiet der möglichen x' von den endlichen Grenzen $\pm \frac{1}{2}\pi$ eingeschlossen sein. Mit Rücksicht hierauf wollen wir die Voraussetzug machen, daß — nötigenfalls durch eine passende Transformation — das Intervall der möglichen Argumentwerte stets auf endliche Grenzen gebracht worden sei: man erreicht dadurch in den späteren Entwicklungen eine gewisse Vereinfachung.

Die möglichen Argumentwerte eines K.-G. können entweder eine stetige Mannigfaltigkeit bilden oder aber auf ein System diskreter Zahlen beschränkt sein. So ist z. B. in der oben betrachteten Rekrutentafel das Argument an sich stetig, wenn auch die nur auf volle Zentimeter notierten Zahlenwerte der beobachteten Körperlängen als eine diskrete Mannigfaltigkeit erscheinen. Teilt man dagegen beobachtete Geburtenmengen in Gruppen zu je hundert ab und wählt als Argument x die gruppenweise ermittelte rH. der Knabengeburten, so ist x von vornherein auf die Reihe der Größen 0.00, 0.01, … 0.99, 1.00 beschränkt. Dementsprechend werden wir fortan *stetige* und *unstetige* K.-G. unterscheiden. Der gemischte Fall, wo ein K.-G. nach dem Argument x zugleich stetig und unstetig ist, soll hier nicht weiter

7*

berücksichtigt werden; er besitzt, soweit ich sehen kann, vor der Hand kein praktisches Interesse und würde sich im übrigen nach den weiterhin folgenden Entwicklungen ohne Schwierigkeit mit erledigen lassen.

§ 73. Bei der nun folgenden Erörterung trennen wir die stetigen K.-G. von den unstetigen und beschäftigen uns zunächst mit den erstgenannten, zu denen im besonderen alle Kollektivreihen gehören, deren Argument durch eigentliche Messung ermittelt wird. Hierbei ist vorab die Rolle klar zu legen, die die sogenannte *Abrundung* der in der Urliste oder der Verteilungstafel notierten Argumente spielt.

Die genaue Angabe der Werte einer stetigen Veränderlichen erfordert, wenn mit Dezimalbrüchen gerechnet wird, im allgemeinen unendlich viele Dezimalen, denn ein gewöhnlicher Bruch liefert nur dann einen endlichen Dezimalbruch, wenn die Zerlegung des Nenners in Primfaktoren ausschließlich auf die Zahlen 2 und 5 führt. Da man nun mit unendlich vielen Dezimalen nicht wirklich rechnen kann, so wirft man alle Ziffern hinter einer gewissen, jedesmal festzusetzenden Stelle ab und operiert mit den solchergestalt *abgerundeten* Beträgen. Infolgedessen stellt ein notierter Betrag nicht nur den Wert der wirklich hingeschriebenen Zahl, sondern zugleich auch diejenigen benachbarten Werte dar, die bei der vorgeschriebenen Abrundung in den notierten Betrag übergehen. Werden z. B. in einer Tafel von Körperlängen alle Angaben mit vollen Zentimetern notiert, so steht die Zahl 171 cm für alle Werte, die zwischen 170.5 und 171.5 cm liegen; wird dagegen von vornherein mit vollen Doppelzentimetern gerechnet, so steht die Zahl 170 cm für alle Werte zwischen 169 und 171 cm.

Geometrisch ausgedrückt führen die vorstehenden Bemerkungen zu folgenden Sätzen. Ein in der Urliste notiertes Argument a vertritt alle Abszissenpunkte x, die einer gewissen, den Punkt a einschließenden *Teilstrecke* angehören. Infolgedessen zerfällt nach Maßgabe der vorgeschriebenen Abrundung die ganze Abszissenachse in eine Folge von aneinander stoßenden Teilstrecken, und zwar derart, daß innerhalb einer Strecke die Notierung für die darin enthaltenen Punkte x konstant bleibt, während beim Übergange in eine benachbarte Strecke der notierte Betrag sprungweise wechselt. Mit Rücksicht hierauf wollen wir die Endpunkte der Teilstrecken als *Wechselpunkte* bezeichnen, wobei dann die Abszissenwerte der Wechselpunkte als exakte, von Abrundung freie Angaben aufzufassen sind.

Für die weitere Erörterung machen wir vor der Hand die meistens zutreffende Voraussetzung, daß innerhalb der Urliste durchweg die gleiche Abrundung gebraucht worden sei, daß also die Teilstrecken sämtlich von gleicher Länge und die Wechselpunkte äquidistant seien. Wir können uns dann das System der äquidistanten Wechselpunkte

ohne weiteres, über den Bereich der möglichen Argumente hinaus, nach beiden Seiten hin in dem Gebiete der außermöglichen Argumente fortgesetzt denken. Allerdings stößt man in der Literatur gelegentlich auf beobachtete Reihen, bei denen die Wechselpunkte nicht äqudistant sind: es wird deshalb nachher noch anzugeben sein, wie solche Reihen aufzufassen sind.

Um nun von der gegebenen Urliste aus zu der Verteilungstafel zu gelangen, schlagen wir folgenden Weg ein. Die Urliste enthält die als beobachtet notierten x und läßt unter der gemachten Voraussetzung zugleich die zugrunde gelegte Abrundung erkennen. Daraufhin tragen wir längs der horizontal gedachten Abszissenachse sämtliche Wechselpunkte ein, die bei der vorgeschriebenen Abrundung für die möglichen und außermöglichen Argumente in Betracht kommen. Weiter schreiben wir unterhalb jeder Teilstrecke, die ja von zwei benachbarten Wechselpunkten eingeschlossen ist, die Zahl x hin, die den Punkten der Strecke als gemeinsame abgerundete Notierung zukommt. Endlich schreiben wir über die Teilstrecke die Zahl y, die angibt, wie oft das der Strecke zugehörende x in der Urliste erscheint. Die y sind innerhalb des Gebietes der möglichen x ganze positive Zahlen (mit Einschluß der Null), während die beiden, außerhalb sich anschließenden und bis zu den Stellen $\pm\infty$ reichenden, Gebiete der außermöglichen x beständig den Wert $y = 0$ aufweisen. Hierbei umfaßt dann das System der x teils *leere*, teils *volle* Argumentwerte, je nachdem die y null oder von Null verschieden sind.

Die in der angegebenen Weise hergestellte und längs der ganzen Abszissenachse erstreckte Folge der Zahlenpaare (x, y) ist nun nichts anderes als die primäre Verteilungstafel *Fechners*; man hat nur nötig, die Abszissenachse aufzurichten, so daß die Stelle $-\infty$ nach oben zu liegen kommt.

§ 74. Aus der primären Tafel lassen sich durch Zusammenlegung der Teilstrecken sogenannte *reduzierte* Tafeln herstellen. Um dies zu erläutern, mache ich der Einfachheit halber die zwar nicht notwendige, in der Regel jedoch zutreffende Voraussetzung, daß die notierten x die Abszissen der Streckenmitten angeben. Ferner denken wir uns die Wechselpunkte fortlaufend numeriert und unterdrücken jetzt die Punkte mit ungerader Nummer, so daß sich die ursprünglichen Strecken paarweise zu Doppelstrecken zusammenfügen. Endlich schreiben wir, wenn x', x'' und y', y'' die Werte von x und y für die beiden Komponenten eines Paares bedeuten, an die Doppelstrecke als neues Argument x das arithmetische Mittel $\frac{1}{2}(x' + x'')$ und als Häufigkeitszahl y die Summe $y' + y''$. Dann ist dadurch eine neue Verteilungstafel entstanden, in der die Abrundung doppelt so groß ist, als vorher, während das Mittel $\frac{1}{2}(x' + x'')$ alle Argumentwerte vertritt, die zwischen den Endpunkten der betreffenden Doppelstrecke liegen.

Statt die ursprünglichen Teilstrecken paarweise zu vereinigen, kann man natürlich auch je drei, je vier usw. zusammenfassen und damit stärker reduzierte Tafeln herstellen. *Fechner* behandelt diese Zusammenlegungen sehr ausführlich, weil sie bei ihm eine durchaus wesentliche Vorbereitung für die Untersuchung der K.-G. bilden. Auf dem hier von uns innezuhaltenden Wege haben jedoch die reduzierten Tafeln vorläufig keine besondere Bedeutung, vielmehr genügt es, wenn wir uns an folgende Sätze halten: 1) *die Urliste lehrt das System der anzusetzenden Wechselpunkte kennen,* 2) *die umgeordnete Urliste lehrt die Menge der beobachteten Argumente zwischen je zwei aufeinander folgenden Wechselpunkten kennen,* 3) *die primäre Verteilungstafel stellt die erlangte Kenntnis in tabellarischer Gestalt dar.*

§ 75. Aus der primären Tafel leiten wir jetzt eine Summentafel ab, die die eigentliche Grundlage für die analytische Darstellung des vorgelegten K.-G. bilden wird. Es bezeichne wieder y die Menge der zwischen zwei benachbarten Wechselpunkten beobachteten Argumente, ferner sei Y die Summe der y, die zwischen der Stelle $-\infty$ und einem beliebig herausgegriffenen Wechselpunkte X auftreten, dann gehört zu jedem X ein bestimmtes Y. Läßt man X die ganze Folge der Wechselpunkte durchlaufen, so ist Y anfangs konstant null, geht dann im Gebiete der möglichen X mit ganzzahligen Werten aufwärts, ohne jemals abzunehmen, erreicht am Ausgang des genannten Gebietes die Umfangszahl m und geht von da ab mit dem konstanten Werte m weiter. Schreibt man die sämtlichen Wertepaare (X, Y) untereinander, so erhält man eine Tafel der *absoluten Summenwerte,* in der die erste Spalte die X enthält, während die zweite Spalte die zugehörigen Y angibt. Von dieser Summentafel gelangt man wieder rückwärts zu der primären Verteilungstafel, wenn man in der Y-Spalte jede Zahl von der unmittelbar folgenden abzieht.

Da es bei der Untersuchung eines K.-G. in dem Endergebnis für gewöhnlich nicht auf den genauen Wert des Umfangs m, sondern höchstens auf seine Größenordnung ankommt, so wollen wir, um die besondere Berücksichtigung von m entbehrlich zu machen, statt der Y die Quotienten $Y:m$ einführen und diese mit $\mathfrak{S}(X)$ bezeichnen, so daß Y gleich $m\mathfrak{S}(X)$ wird. Die von X abhängende Größe $\mathfrak{S}(X)$ ist dann offenbar gleich der rH. der unterhalb X beobachteten Argumentwerte. Stellt man die Wertepaare der X und $\mathfrak{S}(X)$ tabellarisch zusammen, so erhält man die Tafel der *relativen Summen,* die wieder in die oben angegebene Tafel der absoluten Summen übergeht, wenn man die aufgeführten $\mathfrak{S}(X)$ mit dem Umfange m multipliziert. Danach kann man also auch von der Tafel der $\mathfrak{S}(X)$ wieder rückwärts zu der primären Verteilungstafel gelangen.

Die Beziehung zwischen den X und $\mathfrak{S}(X)$ stellen wir jetzt geometrisch dar, indem wir zu jeder Abszisse X das entsprechende

$\mathfrak{S}(X)$ als Ordinate abtragen. Dadurch erhalten wir eine Punktreihe, die anfangs in der Abszissenachse verläuft, dann ohne jemals zu sinken im Gebiete der möglichen X zur Ordinate $+1$ aufsteigt und von da ab mit der konstanten Ordinate $+1$ weiter läuft.

§ 76. Die mit den Summenordinaten $\mathfrak{S}(X)$ konstruierte diskrete Reihe der Summenpunkte enthält in geometrischer Einkleidung alles, was nach Maßgabe der primären Verteilungstafel von dem vorgelegten K.-G. ausgesagt werden kann, so daß wir diese Reihe als Grundlage der weiteren Untersuchung nehmen dürfen. Um nun auf stetige Gebilde zu kommen, führen wir neben der Reihe der $\mathfrak{S}(X)$ noch die Funktion $\mathfrak{S}(x)$ der unbeschränkt veränderlichen reellen Größe x ein und setzen für die zu $\mathfrak{S}(x)$ gehörende Kurve folgende Bedingungen fest: 1) die Kurve verläuft stetig und geht durch sämtliche Summenpunkte $\mathfrak{S}(X)$ hindurch, so daß, wenn x mit X zusammenfällt, $\mathfrak{S}(x)$ gleich $\mathfrak{S}(X)$ wird; 2) die Ordinate $\mathfrak{S}(x)$ nimmt niemals ab; 3) die Funktion $\mathfrak{S}(x)$ besitzt eine Ableitung $\mathfrak{V}(x)$, die im allgemeinen stetig ist, also höchstens an einzelnen Stellen unstetig wird.

Die vorstehenden Bedingungen, die in die Punktreihe der $\mathfrak{S}(X)$ eine stetige Folge von Punkten $\mathfrak{S}(x)$ *einschalten*, lassen für den Verlauf von $\mathfrak{S}(x)$ noch eine gewisse Unbestimmtheit übrig, die wir indessen vorläufig auf sich beruhen lassen dürfen, weil sie in der Natur der Sache liegt und sich praktisch als bedeutungslos erweisen wird, sobald man es mit sachgemäß angelegten Beobachtungsreihen zu tun hat. Es hängt das damit zusammen, daß gemessene Größen, auch wenn man von den eigentlichen Messungsfehlern absieht, immer nur innerhalb eines gewissen Abrundungsspielraums bestimmt sind, woraus sich für alle aus den Messungen errechneten Größen ebenfalls ein gewisser Spielraum ergibt.

Die Funktion $\mathfrak{S}(x)$ nennen wir die *Summenfunktion* und die zugehörige Kurve die *Summenkurve* des vorgelegten K.-G. Diese Namen sind bezeichnender, als der von mir bei einer früheren Gelegenheit („Zur Kollektivmaßlehre“ in *Wundts* Philosophischen Studien Band XIV) gebrauchte Ausdruck „Häufigkeitskurve“.

Der Linienzug, welcher in der angegebenen Weise durch Verbindung der beobachteten Summenpunkte $\mathfrak{S}(X)$ entsteht, läßt sich als der beobachtete und deswegen mit dem Einfluß unausgeglichener Zufälligkeiten behaftete Verlauf derjenigen idealen Kurve auffassen, welche man erhalten würde, wenn man die Urliste mit unbegrenzt vielen Gliedern, sowie mit fehler- und abrundungsfrei notierten Argumentwerten aufstellen könnte. Demgemäß erhält man für Größen, die sich durch irgend eine Operation aus der idealen Kurve ergeben, die entsprechenden *beobachteten* Werte, wenn man die Operation an der beobachteten Summenkurve vollzieht. Von solchen Größen interessiert uns hier zunächst die Ableitung $\mathfrak{V}(x)$ der Funktion $\mathfrak{S}(x)$. Da $\mathfrak{S}(x)$

die rH. der Argumente angibt, die bei dem betrachteten K.-G. unterhalb x liegen, so bedeutet das Inkrement $d\mathfrak{S}(x)$ oder $\mathfrak{V}(x)dx$ die rH. der Argumente, welche in das Intervall von x bis $x + dx$ fallen. Denkt man sich dx konstant, so laufen die $\mathfrak{V}(x)$ proportional den Argumentmengen y einer Verteilungstafel, die mit den unendlich kleinen Teilstrecken dx gebildet worden ist. Dem entsprechend bezeichnen wir $\mathfrak{V}(x)$ als die *Verteilungsfunktion* oder auch kurz als die *Verteilung*, die dem betrachteten K.-G. zukommt; die mit den Ordinaten $\mathfrak{V}(x)$ konstruierte Kurve ist dann die *Verteilungskurve*. Liegt für $\mathfrak{V}(x)$ ein analytischer Ausdruck vor, so soll dieser gelegentlich auch als das *Gesetz* der Verteilung bezeichnet werden.

§ 77. Wegen des Verhaltens der Summenfunktion fällt die $\mathfrak{V}$-Kurve für die außermöglichen x in die Abszissenachse und besitzt im Gebiete der möglichen x positive Ordinaten, die an gewissen Stellen auch den Wert Null annehmen können.

Aus der Beziehung $d\mathfrak{S}(x) = \mathfrak{V}(x)dx$ folgt, wenn man zwischen den Grenzen a und b integriert,

$$\int_a^b \mathfrak{V}(t)\,dt = \mathfrak{S}(b) - \mathfrak{S}(a). \tag{1}$$

Die linke Seite ist der Flächeninhalt der Figur, die begrenzt wird von der $\mathfrak{V}$-Kurve, von der Abszissenachse, sowie von den beiden zu $x = a$ und $x = b$ gehörigen Ordinaten; die rechte Seite bedeutet die rH. der Argumente, die zwischen die Grenzen a und b fallen. Hierdurch wird der Zusammenhang zwischen den Argumenten der Summen- und Verteilungsfunktion veranschaulicht. Da

$$\mathfrak{S}(-\infty) = 0, \quad \mathfrak{S}(+\infty) = 1 \tag{2}$$

ist, so erhält man aus (1) noch die besonderen Beziehungen

$$\int_{-\infty}^{x} \mathfrak{V}(t)\,dt = \mathfrak{S}(x), \quad \int_{-\infty}^{+\infty} \mathfrak{V}(t)\,dt = 1. \tag{3}$$

Die von der Verteilungskurve und der Abszissenachse begrenzte Figur besitzt also stets den Flächeninhalt Eins.

Ist $\mathfrak{S}(x)$ oder $\mathfrak{V}(x)$ bekannt, so läßt sich daraus $\mathfrak{V}(x)$ oder $\mathfrak{S}(x)$ herleiten. Ebenso kann man dann auch die primäre oder die reduzierte Verteilungstafel wieder herstellen, sobald das System der Wechselpunkte gegeben ist, denn das über eine Teilstrecke ausgedehnte Integral von $\mathfrak{V}(x)$ liefert die rH. der in diese Strecke fallenden Argumente.

Die Betrachtungen, die vorstehend von der Urliste zu den Funktionen $\mathfrak{S}$ und $\mathfrak{V}$ geführt haben, lassen erkennen, daß man zwei Argumentreihen wohl auseinander zu halten hat, entsprechend den beiden Tafeln, die zu dem gegebenen K.-G. konstruiert werden können. In der Summentafel gehören die darin auftretenden x den Wechselpunkten an: ihnen stehen gegenüber die Werte der Summenfunktion $\mathfrak{S}(x)$

In der Verteilungstafel dagegen, mag es nun die primäre oder eine reduzierte sein, beziehen sich die notierten x auf die Halbierungspunkte der von den Wechselpunkten begrenzten Teilstrecken; ferner sind die den x gegenüber gestellten *Verteilungszahlen* nicht die Werte von $\mathfrak{V}(x)$, sondern die Beträge der über die einzelnen Teilstrecken ausgedehnten Integrale von $\mathfrak{V}(x)$. Wird bei der Konstruktion der Summenpunkte statt der primären Verteilungstafel eine reduzierte zugrunde gelegt, so hat das keine andere Wirkung, als daß in der ursprünglichen Reihe der Summenpunkte gewisse Punkte unterdrückt und entsprechend in der ursprünglichen Summentafel gewisse Zeilen gelöscht werden. Dagegen ändert sich das Aussehen der Verteilungstafel vollständig, sobald man mit ihr eine Reduktion vornimmt. Hiernach ist die Summentafel in analytischer Beziehung das einfachere Gebilde und wird darum auch den natürlichen Ausgangspunkt für die weitere Untersuchung abgeben.

An der Hand der vorstehenden Betrachtungen läßt sich nun noch der oben beiseite geschobene Fall ungleicher Teilstrecken rasch erledigen. Soll die mitgeteilte Urliste oder Verteilungstafel für den Rechner brauchbar sein, so muß sie unzweideutig die Endpunkte der Teilstrecken erkennen lassen, denen die notierten Argumente zuzuweisen sind. Damit sind aber sofort für die Endpunkte der Teilstrecken oder für die Wechselpunkte X die Summengrößen $\mathfrak{S}(X)$ gegeben, und man kann wiederum eine Summentafel aufstellen. Hierbei ist der Umstand, daß die X dieser Tafel nicht mit konstantem Inkrement fortschreiten, vorläufig unwesentlich, wenn er auch, wie sich später zeigen wird, für die rechnerische Bearbeitung des vorgelegten K.-G. gewisse Erschwerungen nach sich zieht.

§ 78. An der Hand der Funktion $\mathfrak{V}(x)$ stellen wir noch die Definition für gewisse, nachher gebrauchte Durchschnittsgrößen auf. Es seien die n willkürlich gewählten Größen $u_1, u_2, \ldots u_n$ gegeben, ferner die n positiven Koeffizienten $a_1, a_2, \ldots a_n$, deren Summe gleich Eins sein soll, dann ist nach der üblichen Ausdrucksweise die Verbindung

$$U = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n \tag{4}$$

das mit den Gewichten $a_1, \ldots a_n$ gebildete Mittel der u. Diese Bezeichnung wollen wir jetzt noch etwas anders fassen. Es seien die a zunächst rationale Zahlen mit dem gemeinsamen Nenner N, so daß die Produkte

$$Na_1 = A_1, \ldots Na_n = A_n$$

ganze Zahlen bedeuten, dann läßt sich (4) in der Gestalt

$$NU = A_1 u_1 + \cdots + A_n u_n \tag{5}$$

schreiben. Hiernach kann man U auch als das gewöhnliche arithmetische Mittel aus einer mit den Zahlen u gebildeten Größenreihe ansehen,

die N Glieder umfaßt, und in der die Zahl u_g mit der Häufigkeit A_g auftritt. Da nun in dieser Größenreihe a_g die rH. des Vorkommens von u_g bedeutet, so dürfen wir weiter sagen, U sei ein arithmetisches Mittel der u, bei dessen Bildung auf die durch die Zahlen a_g gemessene Häufigkeit des Vorkommens der u_g Rücksicht genommen ist.

Die vorstehende Betrachtung bleibt bestehen, wenn die a irrationale Zahlen sind, nur daß dann N über alle Grenzen wächst; sie gilt ebenso auch noch, wenn die Zahl n der unterschiedenen u unendlich wird. Dies vorausgeschickt denken wir uns einen K.-G. mit dem Verteilungsgesetz $\mathfrak{V}(x)$ gegeben, so daß die rH. der Argumente, die dem Intervall von x bis $x+dx$ angehören, durch das Produkt $\mathfrak{V}(x)dx$ bestimmt ist. Ferner sei $T(x)$ eine irgendwie gegebene Funktion von x, deren Werte wir uns für alle x berechnet denken. Dann soll aus diesen Werten das arithmetische Mittel U gebildet werden, und zwar unter Berücksichtigung der Häufigkeit, mit der die x nach Maßgabe des Verteilungsgesetzes $\mathfrak{V}(x)$ auftreten. Zu dem Ende sind in (5) die u durch die Werte von T und die a durch die Produkte $\mathfrak{V}dx$ zu ersetzen, so daß das gesuchte U gleich der über alle reellen x erstreckten Summe der Produkte $T\mathfrak{V}dx$ wird. Dieses Mittel werden wir fortan den *nach der Verteilung* $\mathfrak{V}(x)$ *gebildeten Durchschnitt der Funktion* $T(x)$ nennen und kurz mit $\mathfrak{D}(T(x))$ bezeichnen. Es ist also

$$\mathfrak{D}(T(x)) = \int_{-\infty}^{\infty} T(x)\,\mathfrak{V}(x)\,dx. \tag{6}$$

Da $\mathfrak{V}(x)$ im Gebiet der außermöglichen x beständig verschwindet, so hat man in (6) die Integration tatsächlich nur über das Gebiet der möglichen x auszudehnen.

Wenn die zugrunde gelegte Verteilung, wie es meistens der Fall sein wird, aus dem Gange der Darstellung ersichtlich ist, so werden wir auch kurz nur von dem Durchschnitt der Funktion $T(x)$ sprechen.

§ 79. Mit den vorstehenden Entwicklungen brechen wir einstweilen die Untersuchung der stetigen K.-G. ab und wenden uns zu der Frage, wie sich die gefundenen Sätze für die unstetigen K.-G. gestalten. Die möglichen x bilden also jetzt eine Reihe diskreter Punkte, deren Anzahl in der Regel eine endliche ist. Zu jedem dieser x gehört eine bestimmte Häufigkeitszahl y, die angibt, wie oft das einzelne x in der Urliste auftritt, während die Summe der y den Umfang m des vorgeschriebenen K.-G. liefert. Da jedes x der Urliste nur den hingeschriebenen Zahlenwert, nicht aber auch benachbarte Werte bedeutet, so fallen offenbar jetzt die Teilstrecken und die Wechselpunkte fort.

Setzt man $y = m\mathfrak{U}(x)$, so ist $\mathfrak{U}(x)$ die rH., mit der das Argument x in der primären Verteilungstafel auftritt. Ist ferner $T(x)$ eine be-

liebige Funktion von x, so ist der mit Berücksichtigung jener rH. gebildete Durchschnitt von T durch die Gleichung

$$\mathfrak{D}(T(x)) = \sum T(x)\,\mathfrak{U}(x) \tag{7}$$

gegeben, wo die Summe rechts über alle in der Verteilungstafel vorkommenden x auszudehnen ist. An die Stelle der stetigen Verteilungsfunktion $\mathfrak{V}(x)$ tritt also im vorliegenden Falle die unstetige Wertreihe der $\mathfrak{U}(x)$.

Ist X ein beliebiger Punkt der Abszissenachse, und $\mathfrak{S}(X)$ die Summe der $\mathfrak{U}(x)$, die dem Punkte X vorhergehen, so ist $\mathfrak{S}(X)$ innerhalb der einzelnen Strecken, die von den Stellen $\pm\infty$ und den beobachteten x begrenzt werden, jedesmal konstant, wächst dagegen, wenn X durch einen Punkt x hindurchgeht, sprungweise um den Betrag des entsprechenden $\mathfrak{U}(x)$. An die Stelle der früheren Summenkurve tritt also jetzt eine *Summentreppe* $\mathfrak{S}(x)$, die anfangs in der Abszissenachse liegt, dann zwischen den beiden äußersten beobachteten x treppenförmig unstetig zu der Ordinate Eins aufsteigt, um von da ab mit der konstanten Ordinate Eins weiter zu laufen.

§ 80. Der vorstehend dargelegte Unterschied zwischen den stetigen und den unstetigen K.-G. tritt übrigens bei der praktischen Ausführung stark zurück, indem es nämlich innerhalb gewisser Grenzen möglich ist, einen stetigen K.-G. wie einen unstetigen zu behandeln, und umgekehrt. Wenn z. B. ein unstetiger K.-G. mit der Verteilung $\mathfrak{U}(x)$ vorliegt, so bestimme man zunächst auf der Abszissenachse die Halbierungspunkte X der Strecken zwischen den aufeinander folgenden beobachteten x und füge überdies vor dem ersten und hinter dem letzten x noch je einen Punkt X an beliebiger Stelle hinzu. Weiter trage man in diesen Punkten X und an den Stellen $\pm\infty$ als Ordinaten die Werte der Summenfunktion $\mathfrak{S}$ ab, die aus der Verteilungsfunktion $\mathfrak{U}(x)$ des vorgelegten K.-G. entspringt. Endlich lege man durch die neu entstandene Punktreihe eine stetige Kurve hindurch, und zwar genau in derselben Weise, wie dies oben bei den stetigen K.-G. dargelegt wurde. Dann kann der entstandene Linienzug als die Summenkurve eines stetigen K.-G. angesehen werden, dem eine gewisse stetige Verteilungsfunktion $\mathfrak{V}(x)$ zukommt. Man hat also jetzt zwei K.-G. vor sich, nämlich den vorgelegten unstetigen und den daraus abgeleiteten stetigen, deren Summenkurven in den X-Punkten zusammenfallen, wobei diese Punkte für den stetigen K.-G. offenbar die Rolle von Wechselpunkten spielen.

Geht man umgekehrt von einem stetigen K.-G. und der zugehörigen Summentafel aus, so denke man sich über jedem Wechselpunkte X der Tafel durch den zugehörigen Summenpunkt parallel zur Abszissenachse eine Gerade gezogen, die nach beiden Seiten bis zu den Ordinatenlinien reicht, die den Mitten der beiden dem Punkte

X anliegenden Teilstrecken entsprechen. Dadurch entsteht eine Stufenreihe, die als Summentreppe eines unstetigen K.-G. aufgefaßt werden kann. Die $\mathfrak{U}(x)$ sind dann durch den von Stufe zu Stufe sprungweise erfolgenden Höhenzuwachs gegeben, während als Argumente von $\mathfrak{U}(x)$ die Mitten der Teilstrecken der ursprünglichen Summentafel anzusehen sind.

Selbstverständlich ist die Verwischung des Unterschiedes zwischen den stetigen und unstetigen K.-G. nur so lange angängig, als man sich auf die Betrachtung einzelner Punkte der Summenkurven beschränken darf, denn der Unterschied tritt natürlich mit voller Schärfe hervor, wenn man die Kurven in ihrem ganzen Verlaufe zu verfolgen hat.

Nach den vorstehenden Festsetzungen können wir nunmehr daran gehen, die eingeführten Funktionen analytisch darzustellen.

Elfte Vorlesung.

Das Exponentialgesetz und die Φ-Reihe.

§ 81. Bei der Untersuchung, die uns jetzt zu beschäftigen hat, wollen wir an den Umstand anknüpfen, daß für einzelne K.-G. schon seit langer Zeit eine analytische Darstellung vorhanden ist. Das vornehmste Beispiel hierfür ist das sogenannte *Gaußsche* Gesetz über die Verteilung der Beobachtungsfehler, dessen mathematische Form sich bereits in dem „Gesetz der großen Zahlen“ von *Jakob Bernoulli*[1]) vorgebildet findet. Um den Inhalt der von *Gauß* aufgestellten Formel deutlich zu machen, bedarf es in der Sprache der Kollektivmaßlehre nur weniger Worte. Wird eine unbekannte Größe, z. B. eine Länge, ein Winkel oder dergleichen, wiederholt gemessen, so stimmen wegen der sogenannten unvermeidlichen Beobachtungsfehler die einzelnen Messungen im allgemeinen weder mit dem gesuchten wahren Werte, noch untereinander überein. Bildet man nun aus solchen in genügender Menge angestellten Messungen einen K.-G., dessen Glieder die einzelnen Messungen und dessen Argument die gemessenen Zahlenwerte x bilden, so gehört dazu eine Verteilungskurve, deren Verlauf nach Aussage des *Gaußschen* Gesetzes durch die Gleichung

$$\mathfrak{V}(x) = a \exp\left[-b(x-c)^2\right] \tag{1}$$

gegeben ist, worin von den drei Konstanten a, b, c die beiden ersten wesentlich positiv sind.

---

1) *Jacob Bernoulli*, Ars conjectandi (Opus posthumum). Basileae 1713.

Die Formel (1), die wir auch als *einfaches Exponentialgesetz* oder, wenn kein Mißverständnis zu befürchten ist, nur als „Exponentialgesetz“ (abgekürzt E.-G.) bezeichnen wollen, wurde von *Gauß* in seiner „Theoria motus“ (Abschnitt III) aus rein theoretischen Betrachtungen hergeleitet, die wir nicht weiter zu besprechen brauchen, weil sie für die hier vorliegende Aufgabe ohne Interesse sind. Ebenso können wir den Umstand beiseite lassen, daß der Ausdruck (1), im Gegensatz zu unseren Festsetzungen über die Funktionen $\mathfrak{V}(x)$, den Wert Null erst für unendlich große Argumente annimmt, denn die Abnahme der Exponentialgröße gegen Null hin erfolgt für große x so rasch, daß die entsprechenden Werte von $\mathfrak{V}(x)$ praktisch gleich Null zu setzen sind. Dagegen ist für die Betrachtungen, die uns augenblicklich beschäftigen, der Umstand von Wichtigkeit, daß die aus (1) entspringende Verteilungskurve symmetrisch angeordnet ist, indem die zu $x = c$ gehörige Ordinate nicht nur ein Maximum von $\mathfrak{V}(x)$, sondern auch eine Symmetrieachse der Kurve darstellt. Denn aus dieser Eigenschaft der *Formel* entwickelte sich allmählich ein *Dogma*, nämlich die Vorstellung, daß die Symmetrie der Verteilung eine allgemeine und wesentliche Eigenschaft der Kollektivreihen sei. Das geschah, trotzdem sich dafür — nebenbei bemerkt selbst im Gebiete der Beobachtungsfehler — weder ein logischer noch ein empirischer noch ein praktischer Zwangsgrund beibringen ließ. Es ist zweifellos ein Verdienst *Fechners*, daß er in seiner „Kollektivmaßlehre“ mit der Vorstellung von einer selbstverständlichen Symmetrie der Kollektivreihen aufgeräumt hat. Wie fest jedoch dieses Dogma anfänglich auch bei *Fechner* eingewurzelt war, kann man in seinem Werke zwischen den Zeilen herauslesen: er arbeitet sich erst nach und nach mit einer gewissen Ängstlichkeit und gedrängt durch unabweisbare Tatsachen zu der Einsicht durch, daß die Symmetrie der Verteilungskurven nur ein besonderer, an bestimmte Bedingungen gebundener Grenzfall sei. Selbst ein so scharfer und unbefangener Denker, wie *Fechner*, bedurfte erst eines gewissen moralischen Rucks, um sich zu sagen, daß schlechterdings kein stichhaltiger Grund anzuführen sei, weshalb z. B. die Kurve einer Rekrutentafel symmetrisch sein müsse, und daß ferner eine solche Symmetrie, wenn sie wirklich existierte, geradezu eine im höchsten Grade merkwürdige Erscheinung bilde.

§ 82. Durch die Tatsache, daß es unter den K.-G. Fälle mit greifbar unsymmetrischer Verteilungskurve gibt, wurde *Fechner* gezwungen, an die Stelle der bis dahin fast ausschließlich benutzten Formel (1) einen anderen Ausdruck zu setzen. Der Weg, den er dabei einschlägt, ist folgender. Die in Betracht gezogenen K.-G. besitzen ein Maximum mit Abfall nach beiden Seiten hin. Demgemäß zerschneidet *Fechner* die Verteilungskurve an dem Orte des Maximums,

dem die Abszisse c zukommen möge, in einen vorangehenden und einen nachfolgenden Ast und sagt, daß jeder der beiden Äste einem besonderen Exponentialgesetze folge, dessen größte Ordinate mit dem Maximum der vorgelegten Verteilungskurve zusammenfalle. Es werden demnach für den vorangehenden und den nachfolgenden Ast die beiden Formeln

$$\mathfrak{V}(x) = V'(x) = ah \exp[-h^2(x-c)^2], \quad x < c,$$
$$\mathfrak{V}(x) = V''(x) = bk \exp[-k^2(x-d)^2], \quad x > c$$

angesetzt, in denen wegen der für die Maxima vorgeschriebenen Bedingungen zunächst $c = d$ und $ah = bk$ ist. Eine weitere Bedingung erhält man dadurch, daß der gesamte Flächeninhalt der Verteilungskurve den Wert Eins besitzen muß. Erinnert man sich nun der in der fünften Vorlesung entwickelten Sätze über die Transzendente Φ, die durch die Gleichung

$$\sqrt{\pi}\,\Phi(x) = 2\int_0^x \exp(-t^2)\,dt$$

definiert worden war, so wird, wenn der Index 1 wieder die Ableitung bedeutet,

$$2\,V'(x) = ah\sqrt{\pi}\,\Phi[h(x-c)]_1,$$
$$2\,V''(x) = bk\sqrt{\pi}\,\Phi[k(x-c)]_1.$$

Um den gesamten Flächeninhalt zu erhalten, hat man V' nach x von $-\infty$ bis c, ferner V'' von c bis $+\infty$ zu integrieren und beide Integrale zu vereinigen. Dies gibt, weil die Summe der beiden Integrale gleich Eins sein muß,

$$\begin{aligned} 2 &= a\sqrt{\pi}(\Phi(0) - \Phi(-\infty)) + b\sqrt{\pi}(\Phi(\infty) - \Phi(0)) \\ &= (a+b)\sqrt{\pi}. \end{aligned}$$

Daraus folgt wegen $ah = bk$

$$2k = a(h+k)\sqrt{\pi}, \qquad 2h = b(h+k)\sqrt{\pi},$$

und weiter

$$(h+k)\,V'(x) = hk\,\Phi[h(x-c)]_1,$$
$$(h+k)\,V''(x) = hk\,\Phi[k(x-c)]_1.$$

Das vorstehende *zweiseitige* Gesetz, wie es *Fechner* nennt, enthält also an verfügbaren Parametern die drei Größen c, h, k.

Die sinnreiche Methode, die *Fechner* zur Bestimmung der drei Parameter anwendet, soll hier nicht weiter besprochen werden, desgleichen auch nicht das zweiseitige „logarithmische" Gesetz, das er für gewisse K.-G. einführt. Es ist nämlich leicht zu übersehen, daß die von uns zugrunde gelegte erweiterte Begriffsbestimmung der K.-G., die ja auch willkürlich ersonnene Reihen zuläßt, von vornherein eine

andere analytische Behandlung der Funktionen $\mathfrak{S}$ und $\mathfrak{B}$ verlangt. Denn wenn als Verteilungskurve ein vorläufig beliebig gestalteter Linienzug auftreten kann, so ist nicht daran zu denken, daß sich solche Gebilde allgemein einer einfachen Formel unterwerfen lassen. Aus diesem Grunde sind für unsere Aufgabe auch die von *Pearson* angegebenen Formen (vergl. die Arbeiten von *Pearson* in den *Philosophical Transactions A.* 1895—97) unzureichend, obgleich sie in mathematischer Beziehung eine größere Biegsamkeit besitzen, als das zweiseitige Gesetz *Fechners*.

Um die Fragestellung, auf die es hier ankommt, dentlich zu machen, soll zunächst ein Beispiel herangezogen werden, das einem ganz anderen Gebiete entnommen ist.

§ 83. Wenn ein musikalischer Ton von einem bestimmten Instrument erregt wird, so vollführt jedes Teilchen der umgebenden Luft gewisse periodische Schwingungen. Man denke sich die Ausweichung des Teilchens aus seiner Ruhelage in drei zueinander senkrechte Komponenten zerlegt, ferner sei t die Zeit, c eine gewisse für den Ton charakteristische Konstante und $v = ct$, dann ist, wie man weiß, jede der drei Komponenten durch eine *Fouriersche* Reihe von der Gestalt

$$\begin{aligned} &a_0 + a_1 \cos v + a_2 \cos 2v + a_3 \cos 3v + \cdots \\ &\quad + b_1 \sin v + b_2 \sin 2v + b_3 \sin 3v + \cdots \end{aligned}$$

darstellbar, in der die a, b gewisse, für den gegebenen Ton konstante Koeffizienten bedeuten. Durch diese Koeffizienten wird der Ton nach Stärke und Klangfarbe vollständig charakterisiert, während die von der Schwingungsdauer abhängende Tonhöhe durch die Größe c bestimmt ist. Die Anzahl der von Null verschiedenen a, b ist, streng genommen und von besonderen Fällen abgesehen, unendlich groß, jedoch kommt bei der Darstellung eines beobachteten Tones nur eine begrenzte Anzahl von Koeffizienten als merklich in Betracht, so daß die übrigen ohne Nachteil gleich Null gesetzt werden dürfen.

Das Wesentliche an dem vorstehenden Beispiel besteht offenbar darin, daß eine beobachtete Kurve in einer die Beobachtungen erschöpfenden Weise dargestellt wird durch eine endliche, aber im voraus nicht feststehende Anzahl von Gliedern einer gewissen Reihenentwicklung, und daß dabei der Anschluß zwischen Beobachtung und Rechnung bewirkt wird durch die passende Bestimmung einer Anzahl verfügbarer Konstanten oder „Parameter".

Will man den Grundgedanken, der soeben an dem Beispiel eines periodischen Vorganges erläutert wurde, auf die K.-G. übertragen, so kommt folgendes in Betracht. Die Funktionen und Kurven, mit denen man es in der Kollektivmaßlehre zu tun hat, sind zwar gewissen Grenz- und Stetigkeitsbedingungen unterworfen, im übrigen aber vor der Hand als willkürlich gegeben anzusehen: es handelt sich also,

allgemein gesprochen, zunächst um eine Reihenentwicklung für *willkürliche* Funktionen. Diese Aufgabe läßt nun für den vorliegenden Fall, wie man ohne weiteres aus vorhandenen Beispielen schließen darf, unendlich viele Arten der Lösung zu, die allerdings, wie man ebenfalls aus Beispielen in anderen Gebieten entnehmen kann, für die analytische Darstellung der K.-G. nicht alle gleich zweckmäßig sein werden; es kommt also darauf an, aus der unbegrenzten Mannigfaltigkeit der möglichen Lösungen solche herauszusuchen, die sich den Eigenschaften der am häufigsten auftretenden K.-G. gewissermaßen von selbst und in ähnlicher Weise anschließen, wie das bei periodischen Vorgängen mit der *Fourierschen* Reihe der Fall ist. Endlich wird man noch zu fordern haben, daß die herausgesuchte Lösung auch für die numerische Rechnung hinlänglich bequem sei, denn die angesetzte Reihenentwicklung soll nicht nur ein theoretisches Bedürfnis befriedigen, sondern auch eine ausgiebige Anwendung auf beobachtete K.-G. zulassen.

§ 84. Nachstehend wird nun eine Lösung angegeben werden, die recht weit gehenden Ansprüchen genügt und darauf beruht, daß man an dem Symbol $\mathrm{sg}\,(y-x)$ die $\mathfrak{D}$-Operation vollzieht. Zunächst mögen der besseren Übersicht halber die Vorbereitungen zur Herleitung der gesuchten Lösung kurz zusammengestellt werden.

Das $\mathfrak{D}$-Zeichen, mit dem wir im folgenden fortwährend arbeiten werden, bedeutet einen mit gewissen Gewichten gebildeten Mittelwert aus einer endlichen oder unendlichen Größenreihe. Daraus fließen für das Rechnen mit dem $\mathfrak{D}$-Zeichen einige einfache Regeln, die man sich zweckmäßigerweise ein für allemal einzuprägen hat.

Bedeutet A eine Konstante, so ist $\mathfrak{D}(A) = A$. Ferner ist für eine beliebige Funktion $T(x)$

$$\mathfrak{D}(A\,T) = A\,\mathfrak{D}(T),$$

d. h. konstante Faktoren dürfen nach Belieben vor oder hinter das $\mathfrak{D}$-Zeichen genommen werden. Weiter ist, wenn zwei Funktionen $T(x)$ und $U(x)$ gegeben sind,

$$\mathfrak{D}(T \pm U) = \mathfrak{D}(T) \pm \mathfrak{D}(U).$$

Endlich lassen sich, wenn $T(x)$ außer dem Argument x noch einen von x unabhängigen Parameter t enthält, linear nach dem Parameter auszuführende Operationen beliebig vor oder hinter dem $\mathfrak{D}$-Zeichen vollziehen: als Beispiel können dienen die nach t auszuführende Summation, Integration und Differentiation.

Weiter haben wir uns des in § 36 abgeleiteten Satzes zu erinnern, der auf den Gleichungen

$$u = h(x-c), \qquad v = h(y-c), \qquad t^2 + w^2 = 1, \tag{2.a}$$

$$\mathrm{sg}\,(y-x) = \sum_n t^n \Re(u)_n\, \Phi(v)_n - F \tag{2.b}$$

beruht. Hierin ist die nach den Polynomen $\mathfrak{R}_n$ und den Ableitungen Φ_n fortschreitende Reihe als eine genäherte Darstellung von $\operatorname{sg}(y-x)$ aufzufassen, deren Fehler durch die Größe F angegeben wird. Die Größen x, y, c sind auf endliche Werte zu beschränken, ferner sind die Größen h, t, w beständig positiv zu nehmen, wobei h angebbar von Null verschieden bleiben muß, während dem Parameter w beliebig kleine, aber angebbare Werte zu erteilen sind. Von dem Fehler F gilt dann folgender Satz: wenn A und B zwei fest vorgeschriebene und beliebig kleine angebbare Größen bedeuten, so läßt sich eine Zahl G derart bestimmen, daß für die zwischen 0 und G gelegenen Werte von w der Fehler F zwischen den Grenzen $\pm A$ bleibt, so lange $y-x$ außerhalb eines kritischen Intervalls mit den Grenzen $\pm B$ bleibt; für die innerhalb dieses Intervalls fallenden Werte von $y-x$ liegt F sicher zwischen den Grenzen ± 2.

Den vorstehenden Satz benutzen wir jetzt, wenn wir an der Größe $\operatorname{sg}(y-x)$ die $\mathfrak{D}$-Operation vollziehen.

§ 85. Bei einem stetigen K.-G. mit der Verteilung $\mathfrak{V}(x)$ besitzt der nach $\mathfrak{V}(x)$ genommene Durchschnitt einer Funktion $T(x)$ die Gestalt

$$\mathfrak{D}[T(x)] = \int T(x)\mathfrak{V}(x)\,dx, \tag{3}$$

wo die Integration über das Gebiet der möglichen x auszudehnen ist, das nach der früher getroffenen Festsetzung ganz im Endlichen liegt. Setzt man für $T(x)$ den Ausdruck $\operatorname{sg}(y-x)$, wobei y vorläufig als eine Konstante anzusehen ist, so wird

$$\mathfrak{D}[\operatorname{sg}(y-x)] = \int \operatorname{sg}(y-x)\mathfrak{V}(x)\,dx. \tag{4}$$

Führt man die $\mathfrak{D}$-Operation an der Gleichung (2.b) aus, so entsteht die genäherte Gleichung

$$\mathfrak{D}[\operatorname{sg}(y-x)] = \sum_n t^n \mathfrak{D}[\mathfrak{R}(u)_n]\,\Phi(v)_n, \tag{5}$$

deren Fehler die Form $\mathfrak{D}(F)$ besitzt und nach (3) in der Gestalt

$$\mathfrak{D}(F) = \int F\mathfrak{V}(x)\,dx \tag{6}$$

geschrieben werden kann. Um den Betrag von $\mathfrak{D}(F)$ abzuschätzen denken wir uns, was erlaubt ist, in (6) die Integrationsgrenzen nach den Stellen $\pm\infty$ verschoben; ferner führen wir die oben mit A, B, G bezeichneten Größen ein und zerschneiden das Integrationsgebiet an den Stellen

$$f = y - B, \qquad g = y + B,$$

so daß drei Abszissenabschnitte mit den Grenzen $-\infty$, f, g, $+\infty$ entstehen. Dann liegt die Größe $y-x$ bei den auszuführenden Integrationen außerhalb oder innerhalb der kritischen Grenzen $\pm B$, je nachdem die Integrationsveränderliche x einem der beiden äußeren

Abszissenabschnitte oder aber dem inneren angehört. Weiter schreiben wir unter Berücksichtigung der aufgestellten Abschnitte die Gleichung (6) in der Gestalt

$$\mathfrak{D}(F) = \int_{-\infty}^{f} F\mathfrak{V}(x)\,dx + \int_{f}^{g} F\mathfrak{V}(x)\,dx + \int_{g}^{\infty} F\mathfrak{V}(x)\,dx \tag{7}$$

und ersetzen darin, indem wir den in (5) auftretenden Parameter w unterhalb G halten, den Fehler F durch seine in den einzelnen Abschnitten auftretenden oberen Grenzwerte, nämlich A, 2, A. Damit liefert offenbar die rechte Seite von (7) einen zu großen Betrag, d. h. $\mathfrak{D}(F)$ ist sicher kleiner als der Ausdruck

$$\int_{-\infty}^{f} A\mathfrak{V}(x)\,dx + \int_{f}^{g} 2\,\mathfrak{V}(x)\,dx + \int_{g}^{\infty} A\mathfrak{V}(x)\,dx.$$

Die Summe der beiden äußeren Integrale ist kleiner als A, da ja das vollständige von $-\infty$ bis $+\infty$ genommene Integral über $\mathfrak{V}(x)$ den Wert 1 besitzt. Bezeichnet man ferner mit M den größten Wert von $\mathfrak{V}(x)$ innerhalb des mittleren Abschnitts, so ist das mittlere Integral kleiner als

$$\int_{f}^{g} 2\,M\,dx = 2\,M(g-f) = 4\,BM.$$

Hiernach wird $\mathfrak{D}(F)$ numerisch kleiner als $A + 4BM$, d. h. man kann den Fehler der Gleichung (5) durch Einführung eines hinreichend kleinen Wertes von w oder auch von $1-t$ beliebig klein halten, so lange die Verteilung $\mathfrak{V}(x)$ durchweg endlich bleibt. Ferner besteht die Brauchbarkeit von (5) auch noch für Verteilungen, bei denen unendlich große Werte von $\mathfrak{V}(x)$ auftreten, sobald man für den Parameter y vorschreibt, daß sein Abstand von den Unendlichkeitsstellen der Funktion $\mathfrak{V}(x)$ nicht unter eine gewisse beliebig klein vorzuschreibende angebbare Grenze sinken dürfe.

§ 86. Überträgt man die vorstehenden Betrachtungen auf den Fall eines unstetigen K.-G. mit der Verteilung $\mathfrak{U}(x)$, so gelangt man zunächst wieder zu der Reihe (5), nur daß jetzt der Fehlerausdruck die Gestalt

$$\mathfrak{D}(F) = \sum F\mathfrak{U}(x)$$

annimmt, wo die Summation nach x über die Gesamtheit der möglichen x auszudehnen ist. Da bei dieser Summation die x auf feste diskrete Werte beschränkt bleiben, so hat man, um die Brauchbarkeit von (5) zu sichern, nur zweierlei vorzuschreiben: erstens nämlich, daß $1-t$ genügend klein sei, und zweitens, daß der Abstand der Größe y von den Argumenten der Verteilung $\mathfrak{U}(x)$ oberhalb einer gewissen kleinen Zahl B bleibe.

Sind die x, wie es gewöhnlich der Fall ist, in endlicher Menge vorhanden und äquidistant, so wollen wir die y der Einfachheit halber

in die Mitten der von den x abgegrenzten Teilstrecken legen und uns das System dieser Halbierungspunkte nach vorwärts und rückwärts mit konstanten Abständen fortgesetzt denken. Diese Halbierungspunkte können dann nach einer früheren Bemerkung zugleich als die Wechselpunkte einer stetigen Verteilung $\mathfrak{V}(x)$ aufgefaßt werden, die in ihren Wechselpunkten dieselben Summenwerte wie $\mathfrak{U}(x)$ besitzt.

§ 87. Verschiebt man in (3) die Integrationsgrenzen nach $-\infty$ nnd $+\infty$, was wie wir wissen erlaubt ist, so kann man schreiben

$$\mathfrak{D}[\operatorname{sg}(y-x)] = \int_{-\infty}^{y} \operatorname{sg}(y-x)\mathfrak{V}(x)\,dx + \int_{y}^{\infty} \operatorname{sg}(y-x)\mathfrak{V}(x)\,dx.$$

Die rechte Seite nimmt, wenn man den Verlauf der sg-Größen unter den Integralzeichen beachtet, die Gestalt

$$\int_{-\infty}^{y} \mathfrak{V}(x)\,dx - \int_{y}^{\infty} \mathfrak{V}(x)\,dx$$

an und geht, wenn man sich der Bedeutung der Summenfunktion $\mathfrak{S}(x)$ erinnert, in

$$[\mathfrak{S}(y) - \mathfrak{S}(-\infty)] - [\mathfrak{S}(\infty) - \mathfrak{S}(y)] = 2\,\mathfrak{S}(y) - 1$$

über. Daraus entsteht, wenn wir fortab den Fehler $\mathfrak{D}(F)$ beiseite lassen, die Gleichung

$$2\,\mathfrak{S}(y) - 1 = \sum t^q \mathfrak{D}[\mathfrak{R}(u)_q]\,\Phi(v)_q, \tag{8}$$

wo

$$u = h(x-c), \qquad v = h(y-c) \tag{9}$$

ist. Zu derselben Formel gelangt man, wenn man einen unstetigen K.-G. zugrunde legt und die oben für y angegebene Einschränkung berücksichtigt.

Spaltet man die Gesamtheit der Argumente x des betrachteten K.-G. in die beiden Mengen, die unterhalb und oberhalb y liegen, so erhält man für die zugehörigen rHH. die Werte $\mathfrak{S}(y)$ und $1-\mathfrak{S}(y)$, deren Differenz auf den Ausdruck $2\,\mathfrak{S}-1$, also auf die linke Seite von (8) führt. Die Reihe (8) liefert also eine Darstellung für die Größe, die nach einer früheren Bemerkung als der *relative Überschuß* der unterhalb y gelegenen x zu bezeichnen ist.

Bei der weiteren Benutzung der Reihe (8) werden wir der Einfachheit halber $t=1$ setzen und erhalten damit die *Fundamentalformel für die interpolatorische Darstellung willkürlich gegebener Kollektivreihen* in der Gestalt

$$2\,\mathfrak{S}(y) - 1 = \sum_q \mathfrak{D}[\mathfrak{R}(u)_q]\,\Phi(v)_q. \tag{10}$$

Zu der vorgenommenen Vereinfachung ist zu bemerken, daß die Forderung eines angebbar von der Einheit verschiedenen Wertes der Größe t lediglich zur Sicherung der unbedingten Konvergenz der Reihenentwicklung eingeführt worden ist, also nur in betracht kommt, wenn es sich um die Summation der unendlichen Reihe (5) handelt.

Nun kommt die Benutzung der Reihe (10) zur numerischen Berechnung von $\mathfrak{S}(y)$ praktisch nur dann in Frage, wenn man dabei mit einer gewissen *begrenzten* Gliedermenge ausreicht, wobei es für den Augenblick gleichgültig ist, ob diese Menge fünf oder zehn oder etwa hundert Glieder umfaßt. In einem solchen Falle ist aber der Ansatz $t = 1$ ohne weiteres erlaubt. Kommt es dagegen nicht auf die numerische Berechnung von $\mathfrak{S}(y)$, sondern nur darauf an, die Funktion $\mathfrak{S}(y)$ durch das System der als Reihenkoeffizienten auftretenden Durchschnittsgrößen $\mathfrak{D}(\mathfrak{R})$ zu definieren oder zu charakterisieren, so ist der Wert von t überhaupt gleichgültig, denn die Größen $\mathfrak{D}(\mathfrak{R})$ hängen von t nicht ab und sind bereits durch den Verlauf der Funktion $\mathfrak{S}(y)$ oder $\mathfrak{S}(x)$ vollständig bestimmt. Dies wird sich später noch deutlicher herausstellen, wenn wir an die Untersuchung der Größen $\mathfrak{D}(\mathfrak{R})$ gehen.

§ 88. Sollen die Vorteile, die die Reihe (10) in der Kollektivmaßlehre gewähren kann, ausgenutzt werden, so ist es wesentlich, daß die Berechnung der einzelnen Glieder möglichst bequem gemacht wird. Das Hauptterfordernis hierfür ist die rasche Ermittlung der Werte der Φ-Funktionen, d. h. die Aufstellung hinreichend ausführlicher Tafeln für diese Funktionen bis zu einem gewissen Index hin. Demgemäß sind im Anhange vierstellige Tafeln für die Grundfunktion Φ selber und für die Größen

$$\Phi_1, \quad \Phi_2 : 2, \quad \Phi_3 : 4, \quad \Phi_4 : 8, \quad \Phi_5 : 16, \quad \Phi_6 : 32$$

gegeben, womit man in zahlreichen Fällen vollkommen ausreicht. Weiter zu gehen habe ich vorläufig nicht für angezeigt gehalten; es scheint mir zweckmäßig, in dieser Hinsicht erst einmal die Erfahrungen einer längeren praktischen Erprobung der ganzen Methode abzuwarten.

In den genannten Tafeln sind zu den Φ-Größen die Divisoren 2, 4, ... hinzugefügt worden, um die Annehmlichkeit zu erreichen, daß in der Tafel jeder Φ-Funktion das Maximum sich nicht weit von Eins entfernt. Selbstverständlich wird man die aufgeschlagenen Tafelwerte nicht mit den betreffenden Potenzen von 2 multiplizieren, um die entsprechenden Φ-Größen zu erhalten, sondern vielmehr diese Potenzen einfach als Multiplikatoren zu den Koeffizienten der Reihe (10) schlagen.

Über die Herstellung der Tafeln für die Ableitungen von Φ möge noch folgendes bemerkt werden. Die angesetzten Tafelwerte wurden zunächst so scharf, als es die siebenstelligen Logarithmentafeln zulassen, einzeln gerechnet, dann auf sechs Stellen abgerundet und durch Differenzen geprüft. Ferner habe ich zur Kontrolle noch jeden zehnten Tafelwert unabhängig von dieser ersten Rechnung nach einem abweichenden Formelsystem abgeleitet. Danach ist zu hoffen, daß die vorliegenden Tafeln von der strengen Korrektheit nur vereinzelt um eine Einheit der letzten Stelle abweichen.

§ 89. Als Gesamtergebnis der bisherigen Entwicklung und als festen Ausgangspunkt für die Behandlung der Kollektivmaßlehre haben wir jetzt die Reihe (10) vor uns. Die gefundene Reihe liefert uns eine Lösung der früher gestellten Aufgabe, nämlich Auffindung einer allgemeinen analytischen Darstellung der Kollektivgegenstände. Die Zweckmäßigkeit der gewählten Form hängt von der Zahl der jedesmal mitzunehmenden Glieder ab, und es kann natürlich der Fall eintreten, daß die Φ-Reihe wegen der Menge der mitzunehmenden Glieder unzweckmäßig wird, gerade so, wie bei der Darstellung periodischer Vorgänge die *Fouriersche* Reihe praktisch versagen kann. Wir dürfen jedoch diesen Fall einstweilen beiseite lassen, indem wir sagen, daß wir zunächst zuzusehen haben, wie weit man mit der aufgestellten Lösung überhaupt kommt. Stellt sich dabei heraus, daß andere Lösungen nötig sind, so wird man danach suchen.

Geschichtlich mag noch erwähnt werden, daß bereits *Bessel* in der Abhandlung „*Untersuchungen über die Wahrscheinlichkeit der Beobachtungsfehler* (*Astronomische Nachrichten*, *Band* 15)" sich mit der Aufgabe beschäftigt hat, eine analytische Darstellung gegebener Verteilungsfunktionen zu entwickeln; er findet tatsächlich eine Reihe, deren Glieder sich aus Ableitungen der Funktion Φ zusammensetzen, unterläßt es jedoch diesen, für die ganze Aufgabe wesentlichen, Umstand weiter zu verfolgen, da der von ihm benutzte Ansatz die Bildungsweise der Koeffizienten ganz undurchsichtig macht.

Die nächste Arbeit wird nun darin bestehen, die Eigenschaften der Gleichung (10) oder, wie wir fortan kurz sagen wollen, der Φ-Reihe näher zu untersuchen. Hierbei wollen wir zur Abkürzung des Ausdrucks einige Bestimmungen treffen, an denen weiterhin, soweit nicht das Gegenteil bemerkt ist, immer festgehalten werden soll.

Vorgelegt ist ein stetiger oder unstetiger K.-G. mit dem Argument x, der Summenfunktion $\mathfrak{S}(x)$ und der zugehörigen Verteilung $\mathfrak{V}(x)$ oder $\mathfrak{U}(x)$. Mit der positiven Konstante h und der positiven oder negativen Konstante c wird die Größe

$$u = h(x - c)$$

gebildet, die fortan als das *Hilfsargument* bezeichnet werden soll. Zu dem Hilfsargument werden die Polynome $\mathfrak{R}(u)_q$ angesetzt und aus ihnen die nach $\mathfrak{V}(x)$ oder $\mathfrak{U}(x)$ genommenen Durchschnitte

$$\mathfrak{D}[\mathfrak{R}(u)_q] = D(c, h)_q$$

hergeleitet, die wir auch kurz mit D_q bezeichnen werden, wenn es nicht darauf ankommt, die benutzten Parameter c, h besonders ersichtlich zu machen. Dann hat die gefundene Darstellung die Gestalt

$$2\,\mathfrak{S}(x) - 1 = \sum_q D(c, h)_q\, \Phi(u)_q, \qquad (q = 0, 1, \ldots) \qquad (11)$$

wo D_0 immer gleich Eins ist, und Φ_0 die Grundfunktion Φ selber bedeutet.

Das Gebiet der Veränderlichen x in der Gleichung (11) ist durch die Bedingung beschränkt, daß nur endliche Werte von x in Betracht kommen. Außerdem sind unter gewissen Umständen von jenem Gebiete noch weitere Stellen auszuschließen, wenn man der Zulässigkeit der Gleichung (11) sicher sein will. Versteht man unter B eine beliebig kleine aber angebbare Zahl und bezeichnet die von $x - B$ bis $x + B$ reichende Argumentstrecke als die Umgebung von x, so hat man bei einem stetigen K.-G. in (11) die Umgebung derjenigen Stellen auszuschließen, in denen $\mathfrak{B}(x)$ unendlich wird. Ebenso sind bei einem unstetigen K.-G. die Umgebungen der möglichen Argumentwerte von $\mathfrak{U}(x)$ oder mit anderen Worten die Sprungstellen der Summentreppe auszuschließen.

Differentiiert man (11) nach x, so erhält man bei einem stetigen K.-G. für die Verteilungs-Funktion $\mathfrak{B}(x)$ die Darstellung

$$2\mathfrak{B}(x) = \sum h D(c, h)_q \Phi(u)_{q+1}. \tag{12}$$

Bei einem unstetigen K.-G. hat die Berechnung der Ableitung von $\mathfrak{S}(x)$ keinen Sinn, da ja $\mathfrak{U}(x)$, als Funktion der stetigen Veränderlichen x aufgefaßt, im allgemeinen null ist und nur für die möglichen Argumentwerte einen von Null verschiedenen Wert annimmt.

§ 90. Über die Konstanten c und h war bisher nichts näheres festgesetzt worden, abgesehen von der Bedingung, daß h stets einen angebbaren positiven Wert besitzen muß. Es ist nun im allgemeinen vorteilhaft, bei der Ansetzung von c und h solche Werte zu bevorzugen, durch welche man zugleich gewisse Nebenbedingungen zu erfüllen vermag. Hierzu gelangen wir durch folgende Betrachtungen.

Der Koeffizient D_1 wird, da $\mathfrak{N}(u)_1 = -u$ ist, durch die Gleichung

$$D_1 = \mathfrak{D}(-u) = \mathfrak{D}(hc - hx) = hc - h\mathfrak{D}(x)$$

bestimmt. Hierin tritt die Größe $\mathfrak{D}(x)$ auf, die fortan als der *Argumentdurchschnitt* des vorgelegten K.-G. bezeichnet werden soll. Wählt man für den Parameter c den besonderen Wert $c_0 = \mathfrak{D}(x)$, so wird D_1 offenbar null.

Für das Polynom $\mathfrak{N}(u)_2$ besteht die Gleichung

$$4\mathfrak{N}(u)_2 = 2h^2(x - c)^2 - 1,$$

woraus durch die $\mathfrak{D}$-Operation

$$4D_2 = 2h^2\mathfrak{D}[(x - c)^2] - 1$$

folgt. Setzt man für c den ausgezeichneten Wert c_0 und führt die Größen s und h_0 durch die Bedingungen

$$s^2 = \mathfrak{D}[(x - c_0)^2], \quad 1 = 2(h_0 s)^2$$

ein, so wird für das Wertepaar c_0, h_0 der Koeffizient D_2 gleich Null. Die Reihe (11) erscheint dann in der Gestalt

$$2\mathfrak{S}(x) - 1 = \Phi(u) + D_3\Phi(u)_3 + D_4\Phi(u)_4 + \cdots, \qquad (13.a)$$

$$u = h_0(x - c_0), \qquad (13.b)$$

worin die Glieder mit Φ_1 und Φ_2 fehlen. Diese Gestalt der Φ-Reihe soll fortan als die *Normalform* bezeichnet werden.

Die Bildung von s kann noch etwas anders eingekleidet werden. Setzt man nämlich mit den n willkürlich gegebenen Größen $a_1, \ldots a_n$ die Gleichung

$$nA^q = a_1^q + a_2^q + \cdots + a_n^q$$

an, so pflegt man A das mit den q-ten Potenzen gebildete Mittel der a zu nennen und spricht demgemäß für $q = 2, 3, \ldots$ von einem quadratischen, kubischen usw. Mittel. Hiernach erscheint s als das quadratische, unter Berücksichtigung der Häufigkeiten $\mathfrak{V}(x)$ oder $\mathfrak{U}(x)$ gebildete Mittel der Differenzen $x - c_0$, die ihrerseits als die Abweichungen der x von ihrem Durchschnitt c_0 oder $\mathfrak{D}(x)$ anzusehen sind. Man könnte daher s füglich als die „mittlere Abweichung" des vorgelegten K.-G. bezeichnen. Da jedoch dieser an sich ganz passende Ausdruck häufig auch noch für andere Mittelwerte jener Abweichungen gebraucht wird, so wollen wir der Deutlichkeit halber für die Größe s einen besonderen Namen einführen und sie die *Streuung* des betrachteten K.-G. nennen.

Die Streuung wird in der Theorie der Beobachtungsfehler seit langer Zeit unter der Bezeichnung „mittlerer Fehler" benutzt; dieser Name hat jedoch in der allgemeinen Kollektivmaßlehre keinen Sinn und würde leicht zu ganz schiefen Vorstellungen führen. Ferner trägt in der Fehlertheorie die Größe h_0 den Namen „Präzision", welcher Ausdruck hier ebenfalls nicht paßt. Jedoch will ich dafür keine andere Bezeichnung einführen, sondern h_0 lediglich als eine für die Rechnung bequeme Zwischengröße behandeln, wobei noch folgender Umstand in Betracht kommt. Bedeuten nämlich die x die gemessenen Beträge einer physischen Größe, z. B. einer Länge, einer Masse oder dergleichen, so kommt ihnen eine bestimmte *Dimension* zu, und die gleiche Dimension ist den Größen c, c_0, s zuzusprechen, während h und h_0 die reziproke Dimension besitzen, also in der Regel nicht unmittelbar anschaulich zu machen sind.

§ 91. Die Größe s, für die wir häufig auch das unmittelbar verständliche Zeichen $\operatorname{str}(x)$ gebrauchen werden, ist als eine Maßzahl für die Stärke der Ausbreitung der Argumentwerte anzusehen: je größer s ist, desto größer ist im allgemeinen die Abszissenstrecke, über welche hin die Argumente ausgebreitet oder *zerstreut* sind. Um diesen Zusammenhang ziffermäßig zu verfolgen, denke man sich von

der Stelle $x = c_0$ aus nach beiden Seiten das a-fache der Streuung s abgetragen, wobei a eine beliebige positive Zahl bedeutet. Dann wird die Argumentmenge M, welche zwischen den Punkten

$$x = c_0 - as \quad \text{und} \quad x = c_0 + as \tag{14}$$

enthalten ist, durch die Differenz

$$M = \mathfrak{S}(c_0 + as) - \mathfrak{S}(c_0 - as)$$

ausgedrückt. Das Hilfsargument u der Normalform nimmt für die beiden Punkte (14) die Gestalt $\pm v$ an, wo

$$v = h_0 as = a\sqrt{0.5} = 0.7071\, a$$

ist. Daraus folgt für M der Ausdruck

$$M = \Phi(v) + D_4 \Phi(v)_4 + D_6 \Phi(v)_6 + \cdots, \tag{15}$$

den wir in der Gestalt

$$M = A + 8 D_4 B + 32 D_6 C + \cdots \tag{16}$$

schreiben wollen. Die Größen $A, B, C, \ldots$ hängen nur von a ab; eine Vorstellung über ihren Verlauf gibt das nachstehende Täfelchen.

a	A	B	C
0.5	+ 0.383	+ 0.484	− 1.106
1.0	+ 0.683	+ 0.484	− 0.726
1.5	+ 0.866	+ 0.146	+ 0.237
2.0	+ 0.955	− 0.108	+ 0.486
2.5	+ 0.988	− 0.143	+ 0.185
3.0	+ 0.997	− 0.080	− 0.040
3.5	+ 1.000	− 0.028	− 0.065

Wie man aus den vorstehenden Zahlen erkennt, ist die Menge der Argumente, die von c_0 um mehr als $3\,s$ abweichen, äußerst gering, falls nicht etwa die Koeffizienten $D_4, \ldots$ größere Werte annehmen. In dieser Beziehung mag hier bemerkt werden, daß bei der weitaus überwiegenden Zahl der bisher genauer untersuchten Kollektivreihen die Beträge von $8 D_4$ und $32 D_6$ unter 0.1 bleiben. Hiernach kann man die *Faustregel* aussprechen, daß die Hauptmenge der Argumente in einem Abszissenintervall von der sechsfachen Länge der Streuung enthalten ist. Oder auch: das Intervall, welches die Hauptmenge der Argumente enthält, liefert nach Division mit 6 einen genäherten Wert der Streuung. Selbstverständlich ist diese Regel nur für solche K.-G. brauchbar, bei denen die Koeffizienten $D_4, D_6, \ldots$ genügend klein sind. Bemerkenswert ist noch, daß die ungeraden D-Koeffizienten in dem obigen Ausdruck von M herausfallen.

Die Streuung war als das quadratische Mittel aus den Abweichungen $x - c_0$ gefunden worden. Man kann nun fragen, wie sich

dieses Mittel gestaltet, wenn die Abweichungen von einer anderen Stelle c aus gerechnet werden. Setzt man die Identität

$$(x-c)^2 = (x-c_0)^2 + 2(x-c_0)(c_0-c) + (c_0-c)^2$$

an und nimmt davon den Durchschnitt, so fällt der mittlere Term auf der rechten Seite heraus und man erhält

$$\mathfrak{D}[(x-c)^2] = s^2 + (c_0-c)^2.$$

Das quadratische Mittel der Abweichungen wird also ein Minimum, wenn $c = c_0$ ist, und fällt um so größer aus, je weiter sich c von c_0 entfernt. Dieser Umstand läßt erkennen, daß $\mathrm{str}(x)$ unter den quadratischen Mitteln der Abweichungen in der Tat eine besondere Stellung einnimmt und deshalb eine Bevorzugung verdient.

§ 92. Die Φ-Reihe benutzt als Bestimmungsstücke der Summen-Funktion die Größen $\mathfrak{D}(x)$, $\mathrm{str}(x)$ und die mit den normalen Werten der Parameter c, h berechneten Koeffizienten $D(c, h)_q$ von $q = 3$ ab. Diese Größen beruhen in letzter Linie auf den Potenzmitteln der Abweichungen $x - \mathfrak{D}(x)$. Die Verwendung solcher Potenzmittel zur Charakteristik gegebener Verteilungen reicht sehr weit zurück; sie ist in der Tat der nächstliegende und einfachste Weg, um zu einer ziffermäßigen Charakterisierung der Kollektivreihen zu gelangen. Es versteht sich jedoch von selbst, daß noch andere Bestimmungsstücke zugrunde gelegt werden können. Deshalb mögen der Vollständigkeit halber noch die von *Fechner* als *Hauptwerte* bezeichneten Argumente erwähnt werden, nämlich der arithmetische Mittelwert, der Zentralwert und der dichteste Wert.

Der *arithmetische Mittelwert* bei *Fechner* ist identisch mit der oben als Argumentdurchschnitt bezeichneten Größe $\mathfrak{D}(x)$. Der *Zentralwert* ist durch die Bedingung $\mathfrak{S}(x) = 0.5$ gegeben und bezeichnet die Stelle, welche gleich viele x über und unter sich hat. Eine rasche, wenn auch meistens nicht sonderlich scharfe Bestimmung des Zentralwertes erhält man, wenn man die Glieder der Urliste oder die beobachteten Argumente nach ihrer Größe fortlaufend numeriert und dann das Glied mit der mittleren Nummer heraussucht. Der *dichteste Wert* oder das *Dichtigkeitsmittel* gibt das Argument an, das am häufigsten vorkommt, also zu dem größten Werte von $\mathfrak{W}(x)$ oder $\mathfrak{U}(x)$ gehört; in der Umgebung des dichtesten Wertes drängen sich daher die Glieder des K.-G. am stärksten zusammen. Ist die Verteilung symmetrisch und besitzt sie überdies nur ein einziges Maximum, so fallen die drei genannten Hauptwerte zusammen.

Die Ermittlung des zweiten und dritten Hauptwertes kann in einzelnen Fällen als summarische Charakteristik ein Interesse besitzen, ist indessen von untergeordneter Bedeutung, wenn es sich um die Untersuchung der allgemeinen Eigenschaften der K.-G. handelt.

Sind in der Normalform der Φ-Reihe die Koeffizienten $D_3, D_4, \ldots$ klein genug, um je nach der vorgeschriebenen Rechnungsschärfe vernachlässigt werden zu dürfen, so liefert ihre Unterdrückung die Gleichungen

$$u = h_0(x - c_0), \qquad 2\mathfrak{S}(x) - 1 = \Phi(u),$$
$$2\mathfrak{V}(x) = h_0 \Phi(u)_1, \qquad \sqrt{\pi}\,\mathfrak{V}(x) = h_0 \exp.(-u^2),$$

d. h. $\mathfrak{V}(x)$ geht in das einfache Exponentialgesetz über. Hiernach kann man, je nach dem Betrage der unterdrückten Glieder, die K.-G. in zwei große Gruppen spalten: die erste umfaßt alle Formen, bei denen das Exponentialgesetz als eine, wenn auch manchmal nur rohe Annäherung gelten darf, die zweite dagegen enthält die Formen, bei denen das genannte Gesetz nicht einmal als rohe Annäherung anzusehen ist. Durch diese Bemerkung wird die Exponentialformel, die eine Zeit lang als eine Art Naturgesetz aufgefaßt wurde, in das richtige Licht gerückt; sie spielt für eine umfangreiche Gruppe der in der Kollektivmaßlehre zu behandelnden Verteilungen eine ähnliche Rolle, wie etwa in der Erdmessung die Vorstellung eines kugelförmigen Erdkörpers.

Bei dem Exponentialgesetz findet das Maximum von $\mathfrak{V}(x)$ für $x = c_0$ statt und besitzt den Wert $h : \sqrt{\pi}$. Diese Eigenschaft kann man gelegentlich benutzen, um rasch zu einer Schätzung von h_0 und damit von $\operatorname{str}(x)$ zu gelangen.

Setzt man bei dem Exponentialgesetz für x die Werte $c_0 \pm w$ ein, so folgt aus den Gleichungen

$$2\mathfrak{S}(c_0 + w) - 1 = \Phi(h_0 w), \qquad 2\mathfrak{S}(c_0 - w) - 1 = \Phi(-h_0 w)$$

durch Subtraktion

$$\mathfrak{S}(c_0 + w) - \mathfrak{S}(c_0 - w) = \Phi(h_0 w).$$

Wir wollen nun w so wählen, daß das Intervall mit den Grenzen $c_0 \pm w$ gerade die Hälfte der Argumente enthält. Es ist dann zu setzen

$$\Phi(h_0 w) = 0.5,$$

woraus $h_0 w = 0.47694$ und weiter durch Multiplikation mit $s\sqrt{2}$

$$w = 0.6745\, s$$

folgt. Hierfür kann man die meistens ausreichende Abkürzung $w = 2\, s : 3$ setzen.

Die Größe w wird vielfach als die *wahrscheinliche Abweichung* bezeichnet und spielte seither bei den Untersuchungen, die sich mit der Voraussetzung des einfachen Exponentialgesetzes begnügen, eine Rolle als Maß für die Ausbreitung der Argumentwerte; sie wird jedoch als solches wertlos, sobald man die Aufgabe der Kollektivmaßlehre

allgemeiner und schärfer als früher faßt. Setzt man nämlich für einen K.-G. mit beliebiger Verteilung wieder die Gleichung

$$\mathfrak{S}(c + w) - \mathfrak{S}(c - w) = 0.5 \tag{17}$$

an und fügt dazu, weil die zwei Größen c und w zu bestimmen sind, zur Festlegung von c eine passende Nebenbedingung, so wird sich bei einem stetigen K.-G. im allgemeinen eine bestimmte Lösung von (17) ergeben. Dagegen kann bei einem unstetigen K.-G. der Fall eintreten, daß die Gleichung (17) unendlich viele Lösungen zuläßt und daß infolgedessen w unbestimmt wird. Um dies einzusehen, nehme man an, daß zunächst *eine* Lösung gefunden worden sei, daß ferner für diese, was ja vorkommen kann, die Werte von $\mathfrak{S}(c + w)$ und $\mathfrak{S}(c - w)$ gleich den Ordinaten zweier Stufen der Summentreppe des betrachteten K.-G. seien, und daß überdies die gefundenen Werte von $c \pm w$ den Mitten dieser Stufen entsprechen. Dann kann man, selbst wenn c festgehalten wird, w innerhalb gewisser Grenzen variieren, ohne daß die Gleichung (17) aufhört erfüllt zu sein: damit wird aber die Lösung von (17) tatsächlich unbestimmt. Während also das durch str (x) gegebene Ausbreitungsmaß für stetige und unstetige K.-G. gleich gut und jederzeit anwendbar ist, müßte man den Gebrauch der wahrscheinlichen Abweichung, die ebenfalls nur als Ausbreitungsmaß zu dienen bestimmt ist, von vornherein auf stetige K.-G. beschränken. Unter solchen Umständen darf man die wahrscheinliche Abweichung ruhig dahin verweisen, wohin sie jetzt gehört, nämlich in die Sammlung der „historischen Altertümer".

§ 93. Legt man bei der allgemeinen Untersuchung der Kollektivreihen die Darstellung durch die Funktion Φ zugrunde, wie das fortan geschehen soll, so erscheinen die in der Φ-Reihe auftretenden Parameter, nämlich der Argumentdurchschnitt $\mathfrak{D}(x)$, die Streuung str (x) und die zur Normalform gehörigen Werte der D-Koeffizienten jedesmal als die eigentlichen Bestimmungsstücke der vorgelegten Verteilung, während der Verlauf der dazu gehörigen Summenfunktion durch die Φ-Reihe selber gegeben ist. Diese Parameter wollen wir fortan als die *numerischen Elemente* der betrachteten Φ-Reihe bezeichnen.

Bei dem einfachen Exponentialgesetz kommen nur zwei Elemente in Betracht. Schreibt man nämlich die zugehörige Verteilung, wie es gewöhnlich geschieht, in der Gestalt

$$\mathfrak{B}(x) = \tfrac{1}{2} h \Phi(u)_1, \quad u = h(x - c), \tag{18}$$

so gehört dazu die Gleichung

$$2\,\mathfrak{S}(x) - 1 = \Phi(u). \tag{19}$$

Es liegt also, da die Koeffizienten D_1 und D_2 verschwinden, die Normalform vor; gleichzeitig sind aber auch alle übrigen D-Koeffizienten null. Umgekehrt muß eine Verteilung, die dem E.-G. gehorcht, auf die

Gleichungen (18) und (19) führen, sobald ihre Darstellung in die Normalform gebracht wird.

Sieht man von dem Falle des E.-G. ab, so erscheinen die numerischen Elemente zunächst in unendlicher Anzahl, wie das bei der Darstellung einer vorläufig willkürlichen Funktion von vornherein selbstverständlich ist. Dieser Umstand schließt jedoch nicht aus, daß im einzelnen Falle schon eine endliche Anzahl von Elementen ausreichend sein kann, um sämtliche Elemente ihrem Werte nach festzulegen. Bei beobachteten Kollektivreihen tritt dieser Fall sogar immer ein, wie die folgende Überlegung lehrt. Die eingeführten Elemente erweisen sich nämlich, wenn man ihrer Entstehungsweise nachgeht, zunächst als Verbindungen der Durchschnittsgrößen $\mathfrak{D}(x)$, $\mathfrak{D}(x^2)$, $\mathfrak{D}(x^3)$, …, die ihrerseits arithmetische Mittel aus den Argumentpotenzen sind, denn wenn man in der, m Glieder umfassenden, Urliste des vorgelegten K.-G. die m darin notierten einzelnen Argumentwerte der Reihe nach mit $x_1, x_2, \ldots x_m$ bezeichnet, so ist

$$m\,\mathfrak{D}(x^p) = \sum_q (x_q)^p, \qquad (q = 1, 2, \ldots m).$$

Solche Potenzsummen lassen sich aber, wie in der Algebra gezeigt wird, durch m unter ihnen, z. B. durch die m niedrigsten, die zu $p = 1$ bis $p = m$ gehören, ausdrücken, woraus das Gleiche für die Größen $\mathfrak{D}(x^p)$ und die aus ihnen abgeleiteten numerischen Elemente folgt.

Die vorstehend angeführte Art von Abhängigkeit zwischen den numerischen Elementen ist allerdings von geringem Interesse, denn ihre Gestalt hängt ausschließlich von der Umfangszahl m ab und hat mit der besondern Form der vorgelegten Verteilung schlechterdings nichts zu tun. Aus diesem Grunde werden im allgemeinen nur solche Abhängigkeiten Beachtung verdienen, deren Gestalt nicht an den Umfang des K.-G. gebunden ist. Derartige Beziehungen treten auf, wenn für die untersuchte Verteilung ein bestimmtes Gesetz, d. h. ein bestimmter analytischer Ausdruck gegeben ist, der nur eine begrenzte Anzahl von willkürlichen Parametern enthält. Beispiele dazu werden uns später begegnen, wenn es sich um theoretisch konstruierte Verteilungen handelt, wie sie aus dem Urnenschema von *Bernoulli* und *Poisson* entstehen.

Wenn man bei der Aufstellung des Systems der für eine Verteilung charakteristischen Bestimmungsstücke darauf Wert legen wollte, daß diese Größen einen möglichst einfachen Zusammenhang mit den beobachteten Argumentwerten besitzen, so würde eine solche Forderung offenbar bei den Potenzmitteln $\mathfrak{D}(x^p)$ ohne weiteres erfüllt sein. Tatsächlich sind denn auch, wie bereits erwähnt wurde, die Potenzmittel von jeher, seit man sich mit der mathematischen Untersuchung von Verteilungen beschäftigt hat, als Bestimmungsstücke herangezogen worden, wie das z. B. schon die meisten älteren Arbeiten über Be-

gründung der Methode der kleinsten Quadrate lehren. Sobald man jedoch der Sache näher rückt, stellt sich heraus, daß es nicht die Werte der Potenzmittel sind, auf die es schließlich ankommt, sondern vielmehr die Werte gewisser, zweckmäßig gewählter, Verbindungen dieser Mittel. Noch deutlicher wird dies durch ein einfaches Beispiel. Es gibt zahlreiche Kollektivreihen, die sich nicht erheblich von dem einfachen Exponentialgesetz entfernen und hinreichend genau dargestellt werden, wenn man in der Φ-Reihe nur bis zu dem Gliede vierter Ordnung geht. Bei solchen Kollektivreihen nimmt die Gleichung für die Summenfunktion in der Normalform die Gestalt

$$2\mathfrak{S}(x) - 1 = \Phi + D_3\Phi_3 + D_4\Phi_4$$

an. Das erste Glied rechts liefert für die Verteilung eine Kurve, welche die leicht zu übersehende glockenförmige Gestalt des einfachen Exponentialgesetzes besitzt. Dazu kommen dann aus den beiden anderen Gliedern rechts gewisse kleine Abweichungen von dem Exponentialgesetz, deren Verlauf sich, wenn D_3 und D_4 numerisch gegeben sind, an der Hand der Tafeln der Φ-Funktionen sofort übersehen läßt. Eine solche Übersicht würde sich aus den Werten von $\mathfrak{D}(x^3)$ und $\mathfrak{D}(x^4)$ *direkt* garnicht gewinnen lassen. Die Sache liegt hier ähnlich wie in der Erdmessung. Um die Unterschiede zwischen hoch und niedrig auf der Erdoberfläche darzustellen, würde der geometrisch einfachste Weg darin bestehen, daß man für die Erdoberfläche die Distanzen der einzelnen Punkte vom Erdschwerpunkt angibt. In Wirklichkeit geschieht dies aber nicht, sondern man benuzt die Abstände von dem sogenannten Meeresniveau, d. h. von einer Referenzfläche, deren Definition alles andere, als geometrisch einfach ist.

§ 94. Geht man daran, beobachtete K.-G. zu untersuchen, so stößt man auf gewisse Fragen, die ziemlich ständig wiederkehren und die man deshalb ein für allemal im voraus zu erledigen suchen wird. Von solchen Fragen sollen hier die nachstehenden behandelt werden:

1) Transformationen der Argumente,
2) Mischung von Argumenten,
3) Mischung von Verteilungen, Unabhängigkeit der Argumente,
4) Unsicherheit errechneter Elemente,
5) Einfluß der Abrundung.

Bevor wir uns jedoch der Untersuchung der genannten Fragen zuwenden, möge zur Abkürzung der Ausdrucksweise die folgende Bemerkung vorausgeschickt werden. Wenn zwei K.-G. mit den Argumenten x und X gegeben sind, so kann man die zugehörigen Summenfunktionen einfach durch die Symbole $\mathfrak{S}(x)$ und $\mathfrak{S}(X)$ ausdrücken, da ja das hinter dem $\mathfrak{S}$-Zeichen stehende Argument jedesmal unzweideutig angibt, welcher von den beiden K.-G. gemeint ist. Der Buchstabe $\mathfrak{S}$ spielt dabei zunächst die Rolle eines Operationszeichens und

zeigt überdies noch die aus der Operation entspringende Funktion an. Das gleiche Verfahren werden wir bei den Zeichen $\mathfrak{W}$, $\mathfrak{U}$, $\mathfrak{D}$, str. usw. innehalten. Der Vorteil, der daraus für die Übersichtlichkeit aller Entwicklungen entspringt, wird weiterhin zur Genüge hervortreten; er beruht wesentlich darauf, daß man für die einzelnen *Größen* nicht immer besondere Zeichen festzusetzen hat.

Zwölfte Vorlesung.

Transformation der Argumente.

§ 95. Bei der Untersuchung eines K.-G. wird der Beobachter für das Argument x im allgemeinen diejenige Maßbestimmung zugrunde legen, welche ihm für die Ausführung der Arbeit als die bequemste erscheint. Mißt er z. B. Schädelumfänge, so wird er einfach die Längen dieser Umfänge, in Millimetern oder einer anderen Maßeinheit ausgedrückt, zum Argument der Kollektivreihe machen. Nachträglich kann sich nun aber herausstellen, daß es bei der weiteren Bearbeitung der Beobachtungen zweckmäßiger ist, das ursprünglich gewählte Argument durch ein anderes zu ersetzen, also z. B. nicht mit den gemessenen Längen, sondern mit deren Logarithmen zu rechnen. In solchen Fällen handelt es sich dann darum, das ursprüngliche Argument zu transformieren und die entsprechende Umwandlung der für den untersuchten K.-G. charakteristischen Größen, d. h. also der numerischen Elemente aufzusuchen.

Soll an die Stelle des ursprünglichen Argumentes x das neue Argument X treten, so muß jedem x, das beobachtet wird oder beobachtet werden könnte, ein bestimmtes X entsprechen und umgekehrt. Wir setzen deshalb $X = f(x)$ und nehmen mit Rücksicht auf die Umstände, die praktisch allein in Betracht kommen, an, daß im Gebiete der möglichen x die Ableitung von $f(x)$ endlich und stetig, sowie durchweg von Null verschieden und positiv sei. Die beiden Veränderlichen x und X steigen und fallen dann gleichzeitig, ferner gehört zu jedem x nur ein X und umgekehrt. Daraus folgt dann noch, daß bei dem vorgelegten K.-G. simultane Werte von x und X immer gleich oft auftreten.

Ist der K.-G. unstetig, so wird, wie die letzte Bemerkung lehrt, $\mathfrak{U}(X) = \mathfrak{U}(x)$. Ist dagegen die Verteilung stetig, und $Y = f(y)$, so ist die Menge der unterhalb Y liegenden X gleich der Menge der unterhalb y liegenden x, d. h. man hat $\mathfrak{S}(Y) = \mathfrak{S}(y)$, woraus durch Differentiieren

$$\mathfrak{W}(Y)\,dY = \mathfrak{W}(y)\,dy$$

folgt. Ist ferner $T(X)$ irgend ein von X abhängender Ausdruck, so wird, je nachdem die Verteilung unstetig oder stetig ist,

$$\mathfrak{D}[T(X)] = \sum T(X)\,\mathfrak{U}(X) = \sum T(X)\,\mathfrak{U}(x),$$

$$\mathfrak{D}[T(X)] = \int T(X)\,\mathfrak{V}(X)\,dX = \int T(X)\,\mathfrak{V}(x)\,dx.$$

Daraus folgt dann, wenn man zusammenfaßt,

$$\mathfrak{D}[T(X)] = \mathfrak{D}[T(f(x))], \tag{1}$$

wo der Durchschnitt links nach X, rechts nach x zu nehmen ist.

Setzt man, um zu den numerischen Elementen von X zu gelangen, für $T(X)$ die Potenz X^{ν}, so hätte man, um die neuen Elemente durch die ursprünglichen auszudrücken, die rechte Seite von (1) analytisch als Funktion der ursprünglichen Elemente darzustellen. Es genügt nun aber, für $f(x)$ einen relativ einfachen Ausdruck, wie z. B. $\operatorname{Log} x$ zu wählen, um sofort zu erkennen, daß man auf diesem Wege für gewöhnlich nicht vorwärts kommen würde. Daraus darf man jedoch nicht schließen, daß die Transformation der Elemente für gewöhnlich überhaupt nicht ernstlich in Betracht zu ziehen sei, denn wir werden später bei der Behandlung numerischer Beispiele sehen, daß die Transformation auf *numerischem* Wege ohne Schwierigkeit und durchaus reinlich ausgeführt werden kann, sobald nur die vorgelegte Verteilungstafel nicht mit übermäßig starker Abrundung behaftet ist. Zu dem Ende hat man, wie schon hier bemerkt werden möge, nur nötig, die Transformation in die Verteilungstafel selber zu verlegen.

§ 96. In einem besonderen Falle verschwinden die Schwierigkeiten der analytischen Transformation, nämlich dann, wenn die Funktion $f(x)$ linear, wenn also

$$X = a(x + b) \tag{2}$$

ist, wo die Größen a, b Konstanten bedeuten. In diesem Falle ist das Ergebnis, wie sich zeigen wird, sehr einfach.

Die angegebene Transformation kommt darauf hinaus, daß man bei der Konstruktion der Summenkurve erstlich den Nullpunkt der Abszissen um die Strecke b zurückschiebt und zweitens die Abszissen mit einem a-mal größeren Maßstabe als vorher abträgt. Dem entsprechend dürfen wir uns die ganze Operation in zwei einfachere und nacheinander auszuführende Transformationen zerlegt denken, die durch die Gleichungen

$$X = x + b \qquad \text{und} \qquad X = ax$$

gegeben sind.

Bezeichnet man die Stücke, die durch die Transformation einander zugeordnet werden, durch die gleichlautenden kleinen und großen

Buchstaben, so kann man für die Transformation $X = x + b$ der Reihe nach ansetzen:

$$X = x + b,$$

$$\mathfrak{D}(x) = c, \quad \mathfrak{D}(X) = C,$$

$$C = \mathfrak{D}(X) = \mathfrak{D}(x + b) = \mathfrak{D}(x) + b = c + b,$$

also weiter

$$C = c + b,$$

$$X - C = x - c.$$

Quadriert man die letzte Gleichung und nimmt dann den Durchschnitt, so wird

$$\operatorname{str}(X)^2 = \operatorname{str}(x)^2,$$

woraus, weil die Streuung immer wesentlich positiv ist,

$$\operatorname{str}(X) = \operatorname{str}(x)$$

folgt. Ferner ergeben sich mit den Gleichungen

$$1 = 2h^2 \operatorname{str}(x)^2, \qquad 1 = 2H^2 \operatorname{str}(X)^2$$

die Beziehungen

$$h = H,$$

$$u = h(x - c), \qquad U = H(X - C),$$

woraus für die Hilfsargumente u, U der zu $\mathfrak{S}(x)$ und $\mathfrak{S}(X)$ gehörigen Normalformen die Gleichungen

$$u = U,$$

$$\mathfrak{R}(u)_p = \mathfrak{R}(U)_p$$

fließen. Damit erhält man für die D-Koeffizienten der Normalform der beiden Φ-Reihen die Ausdrücke

$$D(c, h)_p = \mathfrak{D}[\mathfrak{R}(u)_p] = \mathfrak{D}[\mathfrak{R}(U)_p] = D(C, H)_p,$$

oder

$$D(c, h)_p = D(C, H)_p.$$

Hiernach läßt die Transformation $X = x + b$ die numerischen Elemente ungeändert, mit Ausnahme des Argumentdurchschnitts $\mathfrak{D}(x)$, der sich um den Betrag b verschiebt.

§ 97. Behandelt man in ähnlicher Weise die Transformation $X = ax$ für den Fall eines positiven a, so verläuft die Rechnung folgendermaßen:

$$X = ax,$$

$$C = \mathfrak{D}(X) = \mathfrak{D}(ax) = a\mathfrak{D}(x) = ac,$$

$$X - C = a(x - c).$$

Durch Quadrieren und Mittelnehmen entsteht der Reihe nach

$$\operatorname{str}(X)^2 = a^2 \operatorname{str}(x)^2,$$

$$\operatorname{str}(X) = a \operatorname{str}(x),$$

$$1 = 2H^2 \operatorname{str}(X)^2 = 2h^2 \operatorname{str}(x)^2.$$

$$Ha = h,$$

$$U = H(X - C) = Ha(x - c) = h(x - c) = u,$$

$$\mathfrak{R}(U)_p = \mathfrak{R}(u)_p,$$

$$D(C, H)_p = D(c, h)_p.$$

Die Transformation läßt also wieder die D-Koeffizienten ungeändert, ersetzt aber die Größen $\mathfrak{D}(x)$ und $\operatorname{str}(x)$ durch ihren a-fachen Betrag.

Um nun noch den Fall eines negativen a zu erledigen, genügt es, daß man sich mit den beiden vorhergehenden Transformationen noch die Transformation $X = -x$ kombiniert denkt. Die Rechnung hierzu lautet:

$$X = -x,$$

$$C = \mathfrak{D}(X) = \mathfrak{D}(-x) = -\mathfrak{D}(x) = -c,$$

$$X - C = -(x - c),$$

$$\operatorname{str}(X)^2 = \operatorname{str}(x)^2,$$

$$\operatorname{str}(X) = \operatorname{str}(x), \qquad H = h,$$

$$U = H(X - C) = -h(x - c) = -u.$$

Daraus folgt, weil die $\mathfrak{R}$-Polynome zugleich mit ihrem Index gerade oder ungerade sind,

$$\mathfrak{R}(U)_p = (-1)^p \mathfrak{R}(u)_p,$$

$$D(C, H)_p = (-1)^p D(c, h)_p.$$

Die Transformation $X = -x$ kehrt also das Vorzeichen des Argumentdurchschnitts und der ungeraden D-Koeffizienten um, läßt dagegen die übrigen Elemente, d. h. $\operatorname{str}(x)$ und die geraden D-Koeffizienten ungeändert.

Von den drei betrachteten Transformationen verschiebt die erste die Summenkurve längs der Abszissenachse, während die zweite eine Streckung der Kurve in der Richtung der Abszissen erzeugt. Zu der dritten Transformation ist zu bemerken, daß für $Y = -y$ die rH. der *unterhalb* Y liegenden X oder $\mathfrak{S}(Y)$ gleich der rH. der *oberhalb* y liegenden x, also gleich $1 - \mathfrak{S}(y)$ ist, woraus

$$\mathfrak{S}(Y) + \mathfrak{S}(y) = 1$$

folgt. Man erhält also die transformierte Summenkurve, wenn man der ursprünglichen Kurve erst eine Drehung von 180^0 um den Nullpunkt erteilt und sie dann noch um die Strecke $+1$ in die Höhe schiebt.

§ 98. Die bisher behandelten Transformationen beziehen sich auf das Argument x einer vorgelegten Kollektivreihe. Es sind nunmehr noch die Beziehungen abzuleiten, die auftreten, wenn man das *Hilfsargument* der Φ-Reihe transformiert. Zu dem Ende denken wir uns, daß die Φ-Reihe eines K.-G. in zwei verschiedenen Formen aufgestellt worden sei: einmal mit dem willkürlichen Parameterpaar c, h, sodann aber auch noch mit einem anderen willkürlichen Parameterpaar C, H. Setzt man demgemäß

$$u = h(x-c), \qquad U = H(x-C), \tag{3}$$

$$d_p = D(c, h)_p \;\; = \mathfrak{D}[\mathfrak{R}(u)_p], \tag{4}$$

$$D_p = D(C, H)_p = \mathfrak{D}[\mathfrak{R}(U)_p], \tag{5}$$

so handelt es sich darum, die D_p durch die d_p auszudrücken.

Auf Grund der Reihenentwicklung

$$\exp(-2xv - v^2) = \sum_q \mathfrak{R}(x)_q (2v)^q$$

setzen wir die Ausdrücke

$$t = \sum_p \mathfrak{R}(u)_p (2Hv)^p = \exp(-2uHv - H^2v^2), \tag{6.a}$$

$$T = \sum_q \mathfrak{R}(U)_q (2hv)^q = \exp(-2Uhv - h^2v^2) \tag{6.b}$$

an, aus denen durch Division unter Berücksichtigung der Beziehungen (3)

$$T : t = \exp(-2chHv + 2ChHv - h^2v^2 + H^2v^2)$$

folgt. Mit den Substitutionen $H = gh$ und $hv = w$ erhält man daraus

$$t = \sum_p \mathfrak{R}(u)_p (2gw)^p, \qquad T = \sum_q \mathfrak{R}(U)_q (2w)^q, \tag{7}$$

$$T : t = \exp(-2cHw + 2CHw - w^2 + g^2w^2).$$

Setzt man nun

$$g^2 = 1 - k^2, \quad H(c-C) = fk,$$

$$k^r \mathfrak{R}(f)_r = F_r,$$

so wird zunächst

$$T : t = \exp(-2fkw - k^2w^2) = \sum_r \mathfrak{R}(f)_r (2kw)^r$$

oder

$$T = t \sum_r F_r (2w)^r.$$

Führt man hierin für T und t wieder die Reihen (7) ein, so entsteht die Beziehung

$$\sum_q \mathfrak{R}(U)_q (2w)^q = \left[\sum_p \mathfrak{R}(u)_p (2gw)^p\right] \cdot \left[\sum_r F_r (2w)^r\right].$$

Multipliziert man rechter Hand aus und spaltet dann die Gleichung nach den Potenzen von w, so wird

$$\mathfrak{R}(U)_q = \sum_p \mathfrak{R}(u)_p g^p F_{q-p}, \tag{8}$$

wo die Summe nach p von o bis q zu laufen hat. Führt man endlich noch die $\mathfrak{D}$-Operation aus, so erhält man wegen (4) und (5) die Beziehungen

$$D_q = g^q d_q F_0 + g^{q-1} d_{q-1} F_1 + \cdots + d_0 F_q, \tag{9}$$

deren Anwendung die Kenntnis die Hilfsgrößen

$$g = H : h, \quad k^2 = 1 - g^2, \tag{10}$$

$$f = H(c - C) : k, \tag{11}$$

$$F_r = k^r \mathfrak{R}(f)_r \tag{12}$$

verlangt. Damit ist offenbar die verlangte Transformation hergestellt.

§ 99. Bei der Benutzung von (9) empfiehlt es sich gewöhnlich, den Übergang von (c, h) auf (C, H) nicht in einem Zuge auszuführen, sondern ihn in zwei getrennte Schritte zu zerlegen, indem man z. B. erst von (c, h) auf (C, h) und dann von (C, h) auf (C, H) übergeht.

Für den Übergang von (c, h) auf (C, h) hat man

$$H = h, \quad g = 1, \quad k = 0, \quad f = \infty,$$
$$fk = h(c - C),$$

so daß die F zunächst in der unbestimmten Form $0 \times \infty$ erscheinen. Geht man jedoch auf die in § 33 gegebenen Ausdrücke der $\mathfrak{R}$-Polynome zurück, so wird

$$F_r = [h(C - c)]^r : r!\,.$$

Damit gelangt man zu dem Formelsystem

$$d_p = D(c, h)_p, \quad D_p = D(C, h)_p, \tag{13}$$

$$F_r = [h(C - c)]^r : r!, \tag{14}$$

$$D_q = d_q F_0 + d_{q-1} F_1 + \cdots + d_0 F_q. \tag{15}$$

Wird nur der zweite Parameter h geändert, so ist, wenn man wieder auf die Formeln (9)—(12) zurückgeht, $C = c$ zu setzen. Da dann $f = 0$ wird, so verschwinden nach § 33 die ungeraden F, während für die geraden

$$F_{2r} = (-k^2)^r : 4^r r!$$

ist. Damit gelangt man zu den Formeln

$$d_p = D(c, h)_p, \quad D_p = D(c, H)_p, \tag{16}$$

$$g = H : h, \quad k^2 = 1 - g^2, \tag{17}$$

$$F_{2r} = (-k^2)^r : 4^r r! = (g^2 - 1)^r : 4^r r!, \tag{18}$$

$$D_q = g^q d_q F_0 + g^{q-2} d_{q-2} F_2 + g^{q-4} d_{q-4} F_4 + \cdots. \tag{19}$$

Die gefundenen Formeln finden namentlich dann Verwendung, wenn man bei der Berechnung der numerischen Elemente von der an späterer Stelle zu entwickelnden Summenmethode Gebrauch machen will.

Dreizehnte Vorlesung.

Mischung von Argumenten.

§ 100. Bei der Aufzählung der zu behandelnden Aufgaben war in § 94 nach der Transformation der Argumente die *Argumentmischung* genannt worden. Es soll zunächst der Sinn dieses Ausdrucks an einem Beispiel deutlich gemacht werden.

Gegeben seien n Urnen mit weißen und schwarzen Kugeln. Diese Urnenreihe teilen wir in r Gruppen $G_1, \ldots G_r$, wobei die Gruppe G_k aus n_k Urnen bestehen möge, so daß

$$n = n_1 + n_2 + \cdots + n_r$$

ist. Aus jeder Urne wird unter Zurücklegung der Kugel einmal gezogen. Hierbei notiert man erstlich die Zahl x_k der weißen Kugeln, die aus den Urnen der Gruppe G_k stammen, und zweitens die Gesamtmenge x der gezogenen weißen Kugeln, so daß

$$x = x_1 + x_2 + \cdots + x_r$$

ist. Wird nun dieser Versuch unbegrenzt oft wiederholt, so verteilen sich die beobachteten Werte jedes x_k nach einem gewissen Verteilungsgesetz $\mathfrak{U}(x_k)$, das von der Füllung der Urnen in der Gruppe G_k abhängt. Ebenso verteilen sich die Werte von x nach einem bestimmten Gesetz $\mathfrak{U}(x)$, und man kann dann fragen, wie $\mathfrak{U}(x)$ von den einzelnen $\mathfrak{U}(x_k)$ abhängt.

Will man die vorstehende Aufgabe allgemeiner fassen, so kann das in folgender Weise geschehen. Ein Versuch wird unter konstanten Versuchsbedingungen unbegrenzt oft wiederholt; jeder einzelne Versuch liefert für die r Veränderlichen $x_1, x_2, \ldots x_r$ ein bestimmtes Wertsystem, und zwar derart, daß der beobachtete Wert eines x_k jedesmal völlig unabhängig von den gleichzeitig beobachteten Werten der übrigen x ist; innerhalb der Versuchsreihe folgen die x_k bestimmten Verteilungsgesetzen $\mathfrak{V}(x_k)$ oder $\mathfrak{U}(x_k)$; gesucht wird die Verteilung für die Veränderliche x, die als Funktion der x_k durch einen Ausdruck von der Form

$$x = T(x_1, x_2, \ldots x_r)$$

gegeben ist.

Bei der weiteren Verfolgung der Aufgabe beschränken wir uns auf den Fall, wo x eine mit konstanten Koeffizienten gebildete lineare Verbindung der x_k von der Gestalt

$$x = a_0 + a_1 x_1 + \cdots + a_r x_r \tag{1}$$

ist, und präzisieren die Aufgabe dahin, daß die numerischen Elemente von x durch die Elemente der x_k ausgedrückt werden sollen. Da

man die rechte Seite von (1) durch lineare Transformationen von der Form

$$x_k = A_k X_k + B_k$$

in die Gestalt $X_1 + X_2 + \cdots + X_r$ bringen kann, und da ferner die Elemente der X_k sich in einfacher Weise aus den Elementen der x_k ergeben, so ist es keine wesentliche Einschränkung der Allgemeinheit, wenn wir statt (1) den einfacheren Ausdruck

$$x = x_1 + x_2 + \cdots + x_r \tag{2}$$

zugrunde legen. Wir wollen in diesem Falle sagen: x ist ein *gemischtes Argument*, das durch die *Mischung* der r Argumente x_k entstanden ist.

§ 101. Um bei der Untersuchung die stetigen und unstetigen K.-G. nicht besonders auseinander halten zu müssen, soll eine Vorstellung gebraucht werden, die auch späterhin noch von Nutzen sein wird. Wenn ein Argument x mit seiner Verteilung vorliegt, so denke man sich das Gebiet der unbeschränkten Veränderlichen x in unendlich viele und unendlich kleine Teilstrecken dx von gleicher Länge zerlegt und die rH. der Argumente, die in die Strecke von x bis $x + dx$ fallen, mit $U(x)$ bezeichnet. Dann ist bei einem stetigen K.-G. $U(x)$ gleich $\mathfrak{V}(x)\,dx$, bei einem unstetigen hingegen gleich $\mathfrak{U}(x)$ oder Null, je nachdem die Teilstrecke dx einen möglichen Argumentwert von $\mathfrak{U}(x)$ enthält oder nicht. Die U-Größen für die in (2) auftretenden Argumente sollen einfach durch $U(x_k)$ und $U(x)$ bezeichnet werden, da das hinzugefügte Argument zur Unterscheidung ausreicht. Hiernach ist für eine Funktion von x_k der nach der Verteilung von x_k genommene Durchschnitt durch das Zeichen

$$\mathfrak{D}[T(x_k)] = \sum T(x_k)\, U(x_k) \tag{3}$$

gegeben, wo die Summe über sämtliche Teilstrecken dx_k auszudehnen ist.

Fragt man jetzt weiter nach der Verteilung, welche nicht dem einzelnen Argument x_k, sondern dem ganzen System der x_k und damit auch dem Argument x zukommt, so erhält man, weil die x_k unabhängig voneinander variieren, aus der Produktregel für zusammengesetzte Ereignisse die gesuchte rH. eines bestimmten Wertsystems der x_k in Gestalt des Produktes

$$P = U(x_1)\, U(x_2) \cdots U(x_r). \tag{4}$$

Bedeutet also T eine Funktion der x_k, so erscheint der nach allen Wertsystemen der x_k genommene Durchschnitt von T in der Form

$$\mathfrak{D}(T) = \sum T P, \tag{5}$$

wo die Summe über alle zulässigen Wertsysteme der x_k auszudehnen ist.

§ 102. Für den in (5) angesetzten Durchschnitt gelten gewisse Regeln, die das Rechnen mit dem Ausdrucke $\mathfrak{D}(T)$ erleichtern. Ist

der Ausdruck T frei von x_1, so kann man das Produkt TP in der Gestalt $U(x_1)W$ schreiben, wo W als Faktoren außer T noch die $U(x_k)$ mit den Nummern $k = 2$ bis $k = r$ enthält. Man kann dann ansetzen

$$\mathfrak{D}(T) = [\sum U(x_1)] \cdot [\sum W],$$

wo rechter Hand die Summation in dem ersten Faktor nur nach x_1, in dem zweiten dagegen nach den übrigen x_k vorzunehmen ist. Nun nimmt der erste Faktor wegen der Bedeutung von $U(x_1)$ den Wert Eins an, so daß sich $\mathfrak{D}(T)$ auf die nach $x_2, \ldots x_r$ genommene Summe von W reduziert. Hiernach darf man, wenn T von x_1 frei ist, in dem Produkt TP den Faktor $U(x_1)$ einfach unterdrücken und die Summation von TP so ausführen, als wenn das Argument x_1 überhaupt nicht vorhanden wäre. Der gleiche Schluß gilt, wenn in T mehrere Argumente fehlen, für sämtliche fehlenden Argumente.

Läßt sich T in zwei Faktoren R und S derart zerlegen, daß die einzelnen x_k immer nur in dem einem oder dem andern der beiden Faktoren vorkommen, so denke man sich die $U(x_k)$, deren Argumente in R vorkommen, zu dem Produkte M vereinigt und das Produkt der übrigen $U(x_k)$ mit N bezeichnet. Dann kann man zunächst

$$\mathfrak{D}(T) = [\sum RM] \cdot [\sum SN]$$

schreiben, wo rechter Hand innerhalb der Klammern die Summationen immer nur nach denjenigen x_k auszuführen sind, die hinter dem Summenzeichen in R oder S wirklich vorkommen. Daraus ergibt sich die Regel

$$\mathfrak{D}(T) = \mathfrak{D}(R) \cdot \mathfrak{D}(S),$$

wo rechter Hand in den beiden Faktoren die $\mathfrak{D}$-Operation jedesmal nur nach den dahinter vorkommenden x_k zu bewerkstelligen ist.

§ 103. Nach der vorstehenden Einschaltung kehren wir wieder zu der Gleichung

$$x = x_1 + x_2 + \cdots + x_r = \sum_k x_k \tag{6}$$

zurück und bezeichnen mit $c, h, c_1, h_1, \ldots$ die normalen Werte der Parameter c, h der Hilfsargumente, die bei der Darstellung der Verteilungen von $x, x_1, \ldots$ gebraucht werden. Dann kann man unter Berücksichtigung der soeben entwickelten Regeln ansetzen:

$$x = \sum_k x_k, \quad \mathfrak{D}(x) = \sum_k \mathfrak{D}(x_k), \quad c = \sum_k c_k, \tag{7}$$

$$x - c = \sum_k (x_k - c_k).$$

Quadriert man die letzte Gleichung und nimmt den Durchschnitt, so wird

$$\operatorname{str}(x)^2 = \sum_g \sum_k \mathfrak{D}[(x_g - c_g)(x_k - c_k)]. \tag{8}$$

Sind die Nummern g und k voneinander verschieden, so wird

$$\mathfrak{D}[(x_g - c_g)(x_k - c_k)] = [\mathfrak{D}(x_g - c_g)] \cdot [\mathfrak{D}(x_k - c_k)] = 0,$$

da ja $\mathfrak{D}(x_g - c_g) = \mathfrak{D}(x_g) - c_g = 0$ ist. Man hat also in (8) nur die Glieder zu berücksichtigen, in denen $g = k$ ist. Daraus folgt

$$\operatorname{str}(x)^2 = \sum_k \operatorname{str}(x_k)^2 \tag{9}$$

und weiter für die Parameter h

$$h^{-2} = \sum_k (h_k)^{-2}. \tag{10}$$

Setzt man

$$\operatorname{str}(x_k) = t_k \operatorname{str}(x),$$

so ergibt sich

$$t_k h_k = h, \quad \sum_k t_k^2 = 1. \tag{11}$$

Um nunmehr die D-Koeffizienten zu ermitteln setze man

$$u = h(x - c), \quad u_k = h_k(x_k - c_k), \tag{12}$$

$$A = \exp(-2uv - v^2) = \sum_q \mathfrak{R}(u)_q (2v)^q, \tag{13}$$

$$A_k = \exp(-2u_k t_k v - t_k^2 v^2) = \sum_q \mathfrak{R}(u_k)_q (2t_k v)^q, \tag{14}$$

$$B = \mathfrak{D}(A) = \sum_q D(c, h)_q (2v)^q, \tag{15}$$

$$B_k = \mathfrak{D}(A_k) = \sum_q D(c_k, h_k)_q (2t_k v)^q, \tag{16}$$

$$\sum_k u_k t_k = \sum_k h_k t_k (x_k - c_k) = \sum_k h(x_k - c_k) = h(x - c) = u. \tag{17}$$

Bildet man aus den A_k das Produkt $A_1 A_2 \cdots A_r$ und beachtet für den entstehenden Exponentialausdruck die Relation (17) nebst der zweiten Gleichung in (11), so wird

$$A_1 A_2 \cdots = \exp(-2uv - v^2).$$

Hieraus folgt wegen (13)

$$A = A_1 A_2 \cdots A_r.$$

Führt man hieran die $\mathfrak{D}$-Operation aus und beachtet, daß A_k nur das Argument x_k enthält, so wird

$$\mathfrak{D}(A) = \mathfrak{D}(A_1)\mathfrak{D}(A_2) \cdots \mathfrak{D}(A_r),$$

also wegen (15) und (16)

$$B = B_1 B_2 \cdots B_r. \tag{18}$$

Setzt man hierin für die B-Größen die in (15) und (16) angegebenen Reihen ein und multipliziert rechter Hand aus, so ergeben sich durch Spaltung nach den Potenzen von v die Koeffizienten $D(c, h)_q$ als ganze rationale Verbindungen der Koeffizienten $D(c_k, h_k)_q$ und der Hilfsgrößen t_k. Damit ist die gestellte Aufgabe offenbar gelöst, denn

die Größen $\mathfrak{D}(x)$ und str (x), die mit den $D(c, h)_q$ zusammen das System der numerischen Elemente zu dem Argument x bilden, sind bereits durch (7) und (9) gegeben.

Führt man die Entwickelung von (18) für die niedrigsten Glieder wirklich aus, so wird in B zunächst $D_0 = 1$, $D_1 = 0$ und $D_2 = 0$, wie sich vorhersehen ließ. Ferner erhält man für die drei nächsten Koeffizienten

$$D(c, h)_q = \sum_k D(c_k, h_k)_q t_k^q, \qquad (q = 3, 4, 5). \tag{19}$$

Die späteren Koeffizienten werden verwickelter und sollen hier nicht weiter verfolgt werden, zumal da die Bedeutung des gefundenen Satzes vornehmlich in gewissen daraus zu ziehenden Folgerungen liegt, zu denen wir jetzt übergehen wollen.

§ 104. Wenn die Verteilungen $U(x_k)$ sämtlich die Gestalt des einfachen Exponentialgesetzes besitzen, so sind die Koeffizienten $D(c_k, h_k)_q$, da sie der Normalform angehören, für $q = 1, 2, \ldots$ durchweg null, so daß sich die B_k und folgeweise auch B auf den Wert Eins reduzieren. Daraus folgt, daß sich auch die D-Koeffizienten für x, von D_0 abgesehen, auf Null reduzieren, d. h. auch die Verteilung $U(x)$ besitzt die Gestalt des Exponentialgesetzes. Man kann diese Beziehung passend als den Satz von der *Erhaltung des Exponentialgesetzes* bezeichnen.

Eine andere Folgerung entspringt aus nachstehender Betrachtung. Angenommen die Anzahl r der in der Mischung vereinigten x_k sei sehr groß und beliebiger Steigerung fähig; ferner seien die $U(x_k)$ beliebig beschaffen, mit der einzigen Einschränkung, daß die Streuungen str (x_k) und entsprechend die Größen h_k und t_k untereinander von der gleichen Größenordnung sein sollen. Dann folgt aus (11), daß die t_k von der Ordnung $1 : \sqrt{r}$ sind, da ja die Quadratsumme der t_k gleich Eins sein muß.

Dies festgestellt entwickeln wir $\log B_k$ in die Reihe

$$\log B_k = \sum_k C(k, q)(2 t_k v)^q, \tag{20}$$

wo die $C(k, q)$ ganze rationale Funktionen der $D(c_k, h_k)_q$ sind und die Summation mit $q = 3$ beginnt, weil in B_k die Koeffizienten D_1 und D_2 verschwinden. Summiert man nun die Gleichung (20) nach k, so erhält man links $\log B$, so daß man schreiben kann

$$\log B = \sum_q (2v)^q [\sum_k C(k, q) t_k^q] = \sum_q (2v)^q M(q) T(q),$$

wo

$$M(q) T(q) = \sum_k C(k, q) t_k^q, \quad T(q) = \sum_k t_k^q$$

sein soll. Hierin bedeutet $M(q)$ offenbar einen Mittelwert, der mit den Gewichten $t_1^q, t_2^q, \ldots$ aus den Größen $C(1, q), C(2, q), \ldots$ gebildet worden

ist. Beachtet man nun, daß die t nach r von der Ordnung $-\frac{1}{2}$ sind, so erkannt man, daß $T(q)$ von der Ordnung $-\frac{1}{2}(q-2)$ wird. Die $T(q)$ gehen also, da q mindestens gleich 3 ist, mit wachsendem r gegen Null; das Gleiche gilt von $\log B$, so daß B selber gegen Eins geht, woraus wir schließen, daß das Verteilungsgesetz des gemischten Arguments x dem Exponentialgesetz zustrebt.

Hiernach dürfen wir den Satz aussprechen, daß *der Mischungsprozeß in der Verteilung des gemischten Arguments eine Tendenz nach dem Exponentialgesetz hin erzeugt.* Dieses Ergebnis ist für die Kollektivmaßlehre von wesentlicher Bedeutung, *denn es macht erst verständlich, wie bei zahlreichen Kollektivreihen die tatsächlich vorhandene Annäherung an das Exponentialgesetz überhaupt zustande kommt.*

Der vorstehende Satz findet sich versteckt bei manchen Untersuchungen vor, so z. B. in der Begründung, die *Laplace* für die Methode der kleinsten Quadrate gegeben hat. Bestimmt und klar ausgesprochen wurde der Satz zuerst von *Bessel* in seiner oben (§ 89) genannten Abhandlung über die Wahrscheinlichkeit der Beobachtungsfehler.

Wie weit die aus dem Mischungsprozeß entspringende Tendenz jedesmal durchdringt, hängt natürlich von den Umständen, d. h. von der Beschaffenheit der $U(x_k)$ ab. Man kann z. B. diese Funktionen so wählen, daß $U(x)$ auch für ein unendlich großes r nicht in das Exponentialgesetz übergeht. Zu dem Ende seien die U-Funktionen in unendlicher Anzahl gegeben und so gestaltet, daß die c_k und die D-Koeffizienten für alle U übereinstimmen. Dann sind zunächst die Größen $C(k, q)$ nur von q, aber nicht von k abhängig, so daß wir für $C(k, q)$ kurz $C(q)$ schreiben dürfen und, weil $M(q)$ gleich $C(q)$ wird, die Reihe

$$\log B = \sum_q C(q)\, T(q)\, (2v)^q$$

erhalten. Weiter denken wir uns für die Streuungen die Gleichungen

$$\operatorname{str}(x_k)^2 = 2^{-k}, \quad \operatorname{str}(x)^2 = 2^{-1} + 2^{-2} + \cdots = 1$$

angesetzt, woraus

$$t_k^2 = 2^{-k}, \quad T(q) = 1 : (\sqrt{2^q} - 1)$$

folgt. Die Größen $T(3), T(4), \ldots$ gehen also trotz des unendlichen r nicht gegen Null, und daraus folgt, daß $U(x)$ von dem Exponentialgesetz verschieden ist, sobald es die $U(x_k)$ sind.

Das vorstehende Beispiel läßt deutlich erkennen, weshalb wir oben voraussetzten, daß die $\operatorname{str}(x_k)$ von gleicher Größenordnung sein sollten. Diese Bedingung wird am vollständigsten erfüllt, wenn die $U(x_k)$ übereinstimmende Gestalt besitzen; die Annäherung an das Exponentialgesetz geht dann recht rasch vor sich, wie wir noch an einem Beispiel erläutern wollen.

§ 105. In der Summe

$$N = N_1 + N_2 + \cdots + N_r$$

seien die Größen N_k unabhängig voneinander mit einer gewissen Stellenzahl scharf gerechnet und dann auf eine kleinere, für alle Summanden gleiche, Stellenzahl abgerundet worden. Jedes N_k ist dann mit einem gewissen Abrundungsfehler x_k behaftet und daraus entspringt in N der Fehler

$$x = x_1 + x_2 + \cdots + x_r,$$

der offenbar bis zu dem r-fachen Betrage des einfachen Abrundungsfehlers ansteigen kann. Bei der Aufsuchung von $U(x)$ genügt es, nur x_1 und die davon abhängenden Größen zu betrachten, weil die $U(x_k)$ übereinstimmende Gestalt besitzen. Wählt man nun als Einheit die Einheit der letzten in den N_k beibehaltenen Dezimale, so gehört zu x_1 eine stetige Verteilungsfunktion $\mathfrak{V}(x_1)$, die wegen des bekannten Verhaltens des Abrundungsfehlers x_1 folgende Eigenschaften besitzt. Sie ist außerhalb der durch die Werte ± 0.5 gegebenen Fehlergrenzen beständig null und innerhalb dieser Grenzen konstant, da ja innerhalb der Fehlergrenzen jeder Wert von x_1 gleich leicht vorkommen kann. Ferner ist der konstante Wert von $\mathfrak{V}(x_1)$ gleich Eins, da ja für das zwischen den Fehlergrenzen genommene Integral von $\mathfrak{V}(x_1)$ die Gleichung

$$\int \mathfrak{V}(x_1)\, dx_1 = 1$$

erfüllt sein muß. Da überdies wegen der Symmetrie von $\mathfrak{V}(x_1)$ der Durchschnitt $\mathfrak{D}(x_1)$ verschwindet, so ist der normale Wert des Parameters c_1 null und man erhält darum der Reihe nach

$$\operatorname{str}(x_1)^2 = \mathfrak{D}(x_1^2) = \int_{-0.5}^{0.5} x_1^2\, dx_1 = 1 : 12,$$

$$\operatorname{str}(x_1) = 1 : \sqrt{12}, \quad h_1 = \sqrt{6},$$

$$\operatorname{str}(x) = \sqrt{r} : \sqrt{12}, \quad h = \sqrt{6} : \sqrt{r}, \quad t_1 = 1 : \sqrt{r}.$$

Ferner wird unter Berücksichtigung der Gleichungen (13)—(16)

$$A_1 = \exp(-2hx_1v - t_1^2v^2),$$

$$B_1 = \int_{-0.5}^{0.5} \exp(-2hx_1v - t_1^2v^2)\, dx_1,$$

$$2hvB_1 = \exp(hv - t_1^2v^2) - \exp(-hv - t_1^2v^2),$$

$$B_1 = 1 - \frac{1}{5}(t_1v)^4 + \frac{8}{105}(t_1v)^6 - \cdots.$$

Da nun die B_k einander gleich sind, so wird B gleich der r-ten Potenz von B_1, also

$$B = \sum_q D(c, h)_q (2v)^q = 1 - \frac{v^4}{5r} + \frac{8v^6}{105r^2} + \cdots,$$

und weiter

$$8D(c,h)_4 = -1:10r, \quad 32D(c,h)_6 = 4:105r^2, \quad \text{usw.}$$

Hiernach erhält man z. B. mit $r = 10$ für $8D_4$ den Betrag -0.01. Dies gibt, da das Maximum von $\Phi(u)_4$ gleich 8×0.55 ist, für die Summenfunktion $\mathfrak{S}(x)$ eine Abweichung von dem Exponentialgesetz in dem Betrage

$$\frac{1}{2} \times 0.01 \times 0.55 = 0.0028,$$

d. h. auf 1000 Fälle umgerechnet eine Abweichung von nur 3 Fällen. Noch weniger ergibt das Glied mit D_6.

Das vorstehende Beispiel läßt erkennen, wie rasch sich unter günstigen Umständen der für die $U(x_k)$ charakteristische Verlauf der Verteilung durch den Mischungsprozeß vollständig verwischen kann. Verfolgt man die betrachtete Aufgabe weiter, so gelangt man zu einer ganz hübschen Rechenübung: der Verlauf der Summenkurve läßt sich nämlich aus Bogen von Parabeln r-ter Ordnung zusammensetzen, die an den Stellen, wo sie zusammenstoßen, eine Oskulation von der Ordnung $r-1$ besitzen.

Vierzehnte Vorlesung.

Mischung von Verteilungen. Kriterien der Unabhängigkeit.

§ 106. In der vorhergehenden Vorlesung hat sich gezeigt, daß dem Prozeß der Argumentmischung im allgemeinen die Tendenz nach dem E.-G. hin innewohnt. Außer diesem Mischungsprozeß gibt es nun aber noch einen andern, der als die *Mischung von Verteilungen* bezeichnet werden soll, und der im allgemeinen die gerade entgegengesetzte Tendenz besitzt.

Zum besseren Verständnis möge zunächst an einem Beispiel erläutert werden, worauf der zweite Mischungsprozeß beruht. Für eine gewisse Zeitstrecke und einen gewissen Bezirk von genügendem Umfange seien die Eheschließungen mit Angabe der Altersstufen x und y von Mann und Frau aufgezeichnet worden. Aus einer solchen Masse kann man zunächst eine Kollektivreihe K nebst zugehörigem $\mathfrak{V}(x)$ bilden, indem man die Gesamtheit der Glieder, ohne Rücksicht auf y, lediglich nach dem Argument x ordnet. Andrerseits kann man aber auch die gegebene Masse spalten, indem man jedesmal die Glieder mit einem bestimmt vorgeschriebenen y heraussucht und daraus durch Ordnen nach x eine Teilreihe K_y nebst der zugehörigen Verteilung $\mathfrak{V}(x)_y$ bildet. Denkt man sich auf solche Art erst alle K_y

hergestellt und darauf durch Zusammenwerfen der K_y die Reihe K gebildet, so kann man sagen, daß die Verteilung $\mathfrak{V}(x)$ durch das *Mischen* der vorhandenen $\mathfrak{V}(x)_y$ entstanden sei.

Zunächst wird man nun bei einem K.-G., wie er hier als Beispiel gewählt worden ist, nicht daran denken, die K_y zu „mischen“ und dann nur die aus der Mischung entstandene Reihe K beizubehalten, denn das System der K_y sagt über die vorgelegte Massenerscheinung mehr aus, als die Reihe K, ist also wertvoller. Oft genug ist man jedoch gezwungen, sich mit der Reihe K zu begnügen, weil man das System der K_y entweder garnicht oder doch nur in unbefriedigender Weise aufzustellen vermag. Wenn z. B. die Reihe K nur einen mäßigen Umfang besitzt, so können die Umfänge aller oder doch der meisten K_y so klein ausfallen, daß eine Untersuchung der K_y gegenstandslos wird. Dieser Fall ist sogar als die Regel anzusehen. Kollektivreihen, die 10000 Glieder oder mehr umfassen und dabei die wünschenswerte innere Gleichartigkeit besitzen, verlangen einen ganz erheblichen Arbeitsaufwand und gehören deshalb zu den Seltenheiten. Andrerseits sind bei einem K.-G., der aus natürlichen Vorkommnissen gebildet wird, außer dem einen, als Argument gewählten, Merkmal immer noch unendlich viele andere vorhanden, die sich von Glied zu Glied ändern, also die erstrebte Gleichartigkeit stören und damit den K.-G. tatsächlich zu einem „gemischten“ machen. Hierdurch wird selbstverständlich die Sammlung von wirklich gutem Beobachtungsmaterial, weiter aber auch die richtige Deutung der errechneten Zahlen erschwert.

Durch die vorstehende Bemerkung sieht man sich vor die Aufgabe gestellt zu untersuchen, welche Wirkung der betrachtete Mischungsprozeß im allgemeinen ausüben wird. Hierbei wollen wir uns auf den Fall beschränken, daß nur zwei Argumente zu berücksichtigen sind, da bereits dieser Fall zur Genüge erkennen läßt, wie sich der Gang der Rechnung für mehr als zwei Argumente gestalten würde.

§ 107. Als Ausgangspunkt wählen wir eine Kollektivreihe $K(x, y)$, die von den beiden Argumenten x und y abhängt. Greift man die Glieder heraus, deren y einen vorgeschriebenen Wert besitzt, so erhält man eine Teilreihe $K(x)_y$, deren Argument x ist, während der allen Gliedern gemeinsame Wert von y als der *Index* oder die *Nummer* der Teilreihe bezeichnet werden kann. Ordnet man dagegen die Glieder von $K(x, y)$ nach x, ohne Rücksicht auf den Wert von y, so entsteht die gemischte Reihe $K(x)$, die nur das Argument x enthält. Ebenso entsteht eine gemischte Reihe $K(y)$, wenn man ohne Rücksicht auf x nur nach y ordnet. Die Wirkung der Mischung muß dann in dem Zusammenhang, der zwischen $K(x)$ und den $K(x)_y$ besteht, zum Ausdruck kommen.

Um die auftretenden Beziehungen geometrisch zu veranschaulichen, sollen x und y als rechtwinklige Punktkoordinaten in einer

Grundebene aufgefaßt werden, wobei die positive Richtung nach rechts für die Abszissen x und nach oben für die Ordinaten y gehen möge.

Die Ausbreitung der Punkte, die den möglichen Wertepaaren (xy) entsprechen, kann entweder durchweg stetig oder durchweg unstetig oder endlich teils stetig, teils unstetig sein. Der letzte Fall tritt ein, wenn das eine Argument stetig, das andere unstetig ist. Um diese Fälle nicht immer besonders auseinander halten zu müssen, wenden wir wieder einen schon früher benutzten Kunstgriff an. Durch äquidistante Parallelen zu den Koordinatenachsen werde die Grundebene mit einem Gitter von einander gleichen und unendlich kleinen Quadraten überdeckt. Weiter ordne man jedem Quadrat als *Verteilungsgröße* $U(x, y)$ die rH. derjenigen Glieder von $K(x, y)$ zu, deren Argumentpaar (xy) einen Punkt des betrachteten Quadrats liefert. Endlich bezeichne man die Quadrate als *leer* oder *voll*, je nachdem das zugehörige $U(x, y)$ null oder von Null verschieden ist. Dann stellt der Verlauf von $U(x, y)$ die Verteilung von $K(x, y)$ dar, ohne daß man die stetigen und unstetigen Verteilungen besonders zu scheiden hätte.

§ 108. Bei der weiteren Rechnung möge der an das Summenzeichen Σ angehängte Index x oder y anzeigen, daß die Summation sich über alle Quadrate einer horizontalen oder vertikalen Quadratreihe zu erstrecken habe. Dann gilt zunächst, weil die $U(x, y)$ die rHH. der einzelnen Glieder von $K(x, y)$ bedeuten, die Gleichung

$$\sum_x \sum_y U(x, y) = 1. \tag{1}$$

Ferner geben die Ausdrücke

$$U(x) = \sum_y U(x, y), \qquad U(y) = \sum_x U(x, y) \tag{2}$$

die Verteilung der gemischten Reihen $K(x)$ und $K(y)$ an, denn man erhält die rH., mit der ein bestimmtes Argument x in $K(x)$ auftritt, wenn man alle zu diesem x gehörigen Glieder von $K(x, y)$ ohne Rücksicht auf den Wert von y zusammenfaßt, und das Entsprechende gilt für die Reihe $K(y)$. Aus (2) fließen dann noch die Gleichungen

$$\sum_x U(x) = 1, \qquad \sum_y U(y) = 1. \tag{3}$$

Bezeichnet m den Umfang der Reihe $K(x, y)$, so gibt das Produkt $m\,U(x, y)$ bei festgehaltener Ordinate y an, wie oft die einzelnen x in der Teilreihe $K(x)_y$ auftreten. Da nun der Umfang dieser Teilreihe durch den Ausdruck

$$\sum_x m\,U(x, y) = m\,U(y)$$

gegeben ist, so wird die rH., mit der jedes x in $K(x)_y$ auftritt, gleich

dem Quotienten „$m\,U(x, y)$ dividiert durch $m\,U(y)$“, d. h. die Verteilung des Arguments x in $K(x)_y$ ist durch den Ausdruck

$$U(x)_y = U(x, y) : U(y) \tag{4}$$

bestimmt, dem noch die Beziehung

$$\sum_x U(x)_y = 1 \tag{5}$$

hinzuzufügen ist. Hiermit sind die verschiedenen Verteilungen, die wir zu betrachten haben, festgelegt.

Ist von dem beliebig gewählten Ausdruck $T(x, y)$ der Durchschnitt nach der Verteilung $U(x, y)$ in bezug auf *beide* Argumente zu bilden, so soll dafür das Zeichen $\mathfrak{D}$ ohne weitern Zusatz dienen. Ist ferner der Durchschnitt von T nach der Verteilung $U(x)_y$, also nach dem Argument x bei festgehaltenem y zu nehmen, so soll dies durch das Zeichen $\mathfrak{D}$ mit angehängtem Index x angezeigt werden. Endlich soll das Zeichen $\mathfrak{D}$ mit angehängtem y den Durchschnitt nach der Verteilung $U(y)$ und dem Argument y bedeuten. Setzt man diese Vorschrift in Gleichungen um, so wird

$$\mathfrak{D}(T) = \sum_x \sum_y T \cdot U(x, y), \tag{6}$$

$$\mathfrak{D}_x(T) = \sum_x T \cdot U(x)_y, \tag{7}$$

$$\mathfrak{D}_y(T) = \sum_y T \cdot U(y). \tag{8}$$

Die anderen Durchschnitte, die man noch bilden könnte, sollen beiseite bleiben; sie gehen aus den angegebenen Relationen hervor, wenn man die beiden Argumente ihre Rolle tauschen läßt.

Die Gleichungen (6), (7) und (8) liefern unter Berücksichtigung von (4) noch die Beziehungen

$$\begin{aligned}\mathfrak{D}(T) &= \sum_x \sum_y T \cdot U(y)\, U(x)_y \\ &= \sum_y U(y) [\sum_x T \cdot U(x)_y] \\ &= \sum_y U(y) [\mathfrak{D}_x(T)],\end{aligned}$$

woraus

$$\mathfrak{D}(T) = \mathfrak{D}_y[\mathfrak{D}_x(T)] \tag{9}$$

folgt. Nach dieser Gleichung, die auch bei numerischen Rechnungen sehr wohl als Ausgangspunkt dienen kann, wird zur Bildung von $\mathfrak{D}(T)$ erst nach x gemittelt, und dann von den so entstehenden Durchschnittsgrößen das Mittel nach y genommen.

Reduziert sich T auf einen von x freien Ausdruck, so kann man in

$$\mathfrak{D}(T) = \sum_x \sum_y T \cdot U(x, y) = \sum_x \sum_y T \cdot U(x)_y\, U(y)$$

die Summation nach x ausführen und erhält wegen (5)

$$\mathfrak{D}(T) = \sum_y T \cdot U(y) = \mathfrak{D}_y(T), \tag{10}$$

d. h. den nach $U(y)$ genommenen Durchschnitt der nur von y abhängenden Größe T. Ferner läßt sich, wenn T von y frei ist, in

$$\mathfrak{D}(T) = \sum_x \sum_y T \cdot U(x, y)$$

die Summation nach y ausführen, und man erhält wegen der ersten Gleichung in (2)

$$\mathfrak{D}(T) = \sum_x T \cdot U(x), \tag{11}$$

d. h. den nach $U(x)$ genommenen Durchschnitt der nur von x abhängenden Größe T.

§ 109. Zu den betrachteten Verteilungen gehören gewisse *Summenfunktionen*, die jetzt eingeführt werden sollen. Die Summenfunktionen, die zu den von nur einem Argument abhängenden Reihen

$$K(x), \qquad K(y), \qquad K(x)_y$$

oder den Verteilungen

$$U(x), \qquad U(y), \qquad U(x)_y$$

gehören, bezeichnen wir mit

$$\mathfrak{S}(x), \qquad \mathfrak{S}(y), \qquad \mathfrak{S}(x)_y.$$

Sie bedürfen nach den Entwicklungen der früheren Abschnitte keiner neuen Erläuterung. Neu einzuführen ist dagegen die zur Reihe $K(x, y)$ gehörende Summenfunktion $\mathfrak{S}(x, y)$. Hierbei soll das Zeichen $\mathfrak{S}(X, Y)$ die rH. derjenigen Glieder von $K(x, y)$ bedeuten, die mit ihren Argumenten x, y die beiden Ungleichungen

$$x < X, \qquad y < Y$$

zu gleicher Zeit befriedigen. Dies kommt, geometrisch aufgefaßt, auf folgenden Ansatz hinaus. Man ziehe in der Grundebene von dem Punkte (XY) aus parallel zu den Koordinatenachsen je eine Gerade nach links und nach unten, so daß ein rechter Winkel mit dem Scheitel in (XY) entsteht; ferner summiere man die $U(x, y)$ aller Gitterquadrate, die innerhalb dieses rechten Winkels liegen; dann liefert die gefundene Summe den Wert von $\mathfrak{S}(X, Y)$.

Trägt man den Wert von $\mathfrak{S}(X, Y)$ senkrecht zur Grundebene im Punkte (XY) als dritte Koordinate ab, so erhält man die *Summenfläche* zu $K(x, y)$ deren Verlauf wir uns auf folgende Weise verdeutlichen können. Man suche in der Grundebene zunächst die beiden Gitterquadrate auf, die das kleinste und das größte in $K(x, y)$ vorkommende x besitzen, und ziehe durch diese beiden Stellen je eine Parallele zur y-Achse; ebenso ziehe man durch die beiden Gitterquadrate mit dem kleinsten und dem größten in $K(x, y)$ vorkommenden y je eine Parallele zur x-Achse; dann zerfällt die Grundebene in neun Gebiete

die wir uns in der aus nachstehender Figur ersichtlichen Weise durch die Nummern I, II, ... IX bezeichnet denken.

I	II	III
IV	V	VI
VII	VIII	IX

Das Rechteck V umschließt offenbar alle vollen Quadrate, während die acht übrigen Gebiete nur leere Quadrate enthalten.

Führt man jetzt für beliebige Punkte innerhalb der neun Gebiete die soeben angegebene Konstruktion mit dem rechten Winkel aus, so ergibt sich folgendes. Innerhalb des Gebietes

$$\mathrm{I} + \mathrm{IV} + \mathrm{VII} + \mathrm{VIII} + \mathrm{IX}$$

ist $\mathfrak{S}(X, Y)$ beständig null, d. h. die Summenfläche verläuft in der Grundebene. Dagegen ist innerhalb des Gebietes III der Wert von $\mathfrak{S}(X, Y)$ gleich der Summe aus den $U(x, y)$ aller vollen Quadrate, also beständig gleich Eins und verläuft demnach parallel zur Grundebene in dem konstanten Abstande Eins. Wandert man ferner von dem Gebiete $\mathfrak{S} = 0$ nach dem Gebiete $\mathfrak{S} = 1$ auf einer beliebigen Geraden durch das Gebiet

$$\mathrm{II} + \mathrm{V} + \mathrm{VI}$$

hindurch, so läuft $\mathfrak{S}(X, Y)$ — je nach Umständen stetig oder unstetig — ohne jemals abzunehmen von 0 bis 1.

Wandert der Punkt (XY) parallel zur y-Achse in dem Gebiete II aufwärts, so kommen in der Summe $\mathfrak{S}(X, Y)$ beständig neue $U(x, y)$ hinzu, die jedoch durchweg leeren Quadraten angehören, so daß $\mathfrak{S}(X, Y)$ in Wirklichkeit konstant bleibt. Daher ist die Fläche über II zylindrisch gestaltet und parallel zur y-Achse gerichtet. Ebenso ist die Fläche über VI cylindrisch und parallel zur x-Achse gerichtet. Dagegen ist der Verlauf über V nicht zylindrisch. Man könnte also danach von einem *Summenpodest* sprechen, ähnlich wie wir früher gelegentlich den Ausdruck *Summentreppe* gebraucht haben.

Geht man in II von der Stelle (XY) parallel zur x-Achse um eine Quadratseite vorwärts, so ist die Änderung von $\mathfrak{S}(X, Y)$ gleich der Summe derjenigen $U(X, y)$, deren y unterhalb Y liegt. Diese Summe umfaßt aber *alle* Glieder von $K(x, y)$, für die $x = X$ ist, stimmt also mit $U(X)$ überein. Demnach wächst über dem Gebiete II in dem parallel zur x-Achse genommenen Profil der Summenfläche die Profilhöhe um $U(X)$, wenn man vom Punkte (XY) aus um eine Quadratseite vorwärts geht, d. h. die Profilhöhen entstehen durch Aufsummieren der $U(X)$, und das Profil selber stellt infolgedessen den Verlauf der

Summenkurve $\mathfrak{S}(X)$ dar. In derselben Weise ergibt sich, daß über VI die parallel zur y-Achse genommenen Profile den Verlauf von $\mathfrak{S}(Y)$ darstellen.

In dem Gebiete V hängen die Profile parallel zur x-Achse von den Summenfunktionen $\mathfrak{S}(x)_y$ ab, jedoch ist der Zusammenhang nicht so einfach, daß man ein unmittelbar anschauliches Bild erhielte.

Aus den Summenfunktionen ergeben sich rückwärts noch die Verteilungsfunktionen. Die zu $\mathfrak{S}(x)$, $\mathfrak{S}(y)$, $\mathfrak{S}(x)_y$ gehörigen Verteilungsfunktionen, die wir, je nachdem die vorgelegte Reihe $K(x, y)$ durchweg stetig oder unstetig ist, mit

$$\mathfrak{B}(x), \quad \mathfrak{B}(y), \quad \mathfrak{B}(x)_y$$

oder

$$\mathfrak{U}(x), \quad \mathfrak{U}(y), \quad \mathfrak{U}(x)_y$$

bezeichnen, bedürfen keiner besondern Erläuterung mehr, da es sich um Reihen mit nur einem Argument handelt.

Die Verteilungsfunktion, die zu $\mathfrak{S}(X, Y)$ oder der Summenfläche gehört, wird erhalten, wenn man in dem Quadratgitter nur die vollen Quadrate beachtet und die entsprechenden, von Null verschiedenen Größen $U(x, y)$ zusammenstellt. Diese $U(x, y)$ liefern dann bei einem durchweg unstetigen $K(x, y)$ die Verteilungsfunktion $\mathfrak{U}(x, y)$.

Ist die Reihe $K(x, y)$ durchweg stetig, so schreiben wir die rH. der in ein volles Quadrat fallenden Glieder von $K(x, y)$ in der Gestalt

$$U(x, y) = \mathfrak{B}(x, y)\,dx\,dy,$$

wo dx die horizontale, dy die vertikale Quadratseite bedeutet. Der Ausdruck $\mathfrak{B}(x, y)$ liefert dann die Verteilungsfunktion zu $\mathfrak{S}(x, y)$, und man erhält, da es sich jetzt bei der Bildung von $\mathfrak{S}(X, Y)$ um die Addition von unendlich vielen unendlich kleinen Summenelementen, also um eine Integration handelt,

$$\mathfrak{S}(X, Y) = \iint \mathfrak{B}(x, y)\,dx\,dy,$$

wobei man die Integration nach x von $-\infty$ bis X, nach y von $-\infty$ bis Y zu erstrecken hat. Umgekehrt ergibt sich $\mathfrak{B}(x, y)$, wenn man $\mathfrak{S}(x, y)$ zweimal partiell, erst nach x und dann nach y, differentiiert.

Ist die Reihe $K(x, y)$ teils stetig, teils unstetig, und z. B. x das stetige, y das unstetige Argument, so erhält man zu $\mathfrak{S}(x)$, $\mathfrak{S}(y)$ und $\mathfrak{S}(x)_y$ die Verteilungen $\mathfrak{B}(x)$, $\mathfrak{U}(y)$ und $\mathfrak{B}(x)_y$, wogegen zu $\mathfrak{S}(x, y)$ eine Verteilungsfunktion $\mathfrak{B}(x, y)$ gehört, in der x stetig verläuft, während y auf eine endliche Zahl von diskreten Werten beschränkt bleibt.

§ 110. Nach Einführung der Summenfunktionen hat als nächster Schritt deren analytische Darstellung zu folgen. Zu dem Ende definieren wir den von den zwei Veränderlichen X, x abhängenden Ausdruck $E(X, x)$ durch die Gleichung

$$2E(X, x) = \operatorname{sg}(X - x) + 1. \tag{12}$$

Wegen des Verhaltens der sg-Funktion nimmt $E(X, x)$ den Wert 1 oder $\frac{1}{2}$ oder 0 an, je nachdem $X - x$ positiv oder null oder negativ ist. Damit wird, weil $E(X, x)$ von y frei ist, wegen (11) zunächst

$$\mathfrak{D}[E(X, x)] = \sum_x E(X, x)\, U(x).$$

Spaltet man nun die Summe nach den drei Bedingungen

$$x < X, \qquad x = X, \qquad x > X$$

in drei Teile, so entsprechen diesen Teilen die drei Beiträge

$$\mathfrak{S}(X), \qquad \tfrac{1}{2}U(X), \qquad 0.$$

Der zweite Beitrag ist unendlich klein, wenn die Verteilung von $K(x)$ stetig verläuft; für ein unstetiges $K(x)$ ist er dagegen null, sobald wir wieder wie früher bei unstetigen Kollektivreihen die Festsetzung treffen, daß X niemals mit einem beobachteten x zusammenfallen solle. Demnach dürfen wir setzen

$$\mathfrak{S}(X) = \mathfrak{D}[E(X, x)], \tag{13.a}$$

woraus die uns von früher her bekannte Gleichung

$$2\mathfrak{S}(X) - 1 = \mathfrak{D}[\operatorname{sg}(X - x)] \tag{13.b}$$

folgt. In derselben Weise erhält man für $\mathfrak{S}(Y)$

$$\mathfrak{S}(Y) = \mathfrak{D}[E(Y, y)], \tag{14.a}$$

$$2\mathfrak{S}(Y) - 1 = \mathfrak{D}[\operatorname{sg}(Y - y)]. \tag{14.b}$$

Setzt man weiter den Durchschnitt

$$\mathfrak{D}[E(X, x) E(Y, y)] = \sum_x \sum_y E(X, x) E(Y, y)\, U(x, y)$$

an, so ist die Summation nach x und y in Wirklichkeit nur von $-\infty$ bis X und Y auszuführen, da die übrigen Glieder der Summe verschwinden. Daraus folgt, wenn man sich der Definition von $\mathfrak{S}(X, Y)$ erinnert,

$$\mathfrak{S}(X, Y) = \mathfrak{D}[E(X, x) E(Y, y)] \tag{15.a}$$

oder, wenn man die E-Größen ausmultipliziert,

$$\begin{aligned} 4\mathfrak{S}(X, Y) - 1 = {} & \mathfrak{D}[\operatorname{sg}(X - x)\operatorname{sg}(Y - y)] \\ & + \mathfrak{D}[\operatorname{sg}(X - x)] + \mathfrak{D}[\operatorname{sg}(Y - y)]. \end{aligned} \tag{15.b}$$

Bildet man noch den Ausdruck

$$\Delta[X, Y] = \mathfrak{S}(X, Y) - \mathfrak{S}(X)\mathfrak{S}(Y) \tag{16.a}$$

und führt darin nach (13.b), (14.b) und (15.b) die sg-Größen ein, so entsteht die Gleichung

$$\begin{aligned} 4\Delta[X, Y] = {} & \mathfrak{D}[\operatorname{sg}(X - x)\operatorname{sg}(Y - y)] \\ & - \mathfrak{D}[\operatorname{sg}(X - x)] \cdot \mathfrak{D}[\operatorname{sg}(Y - y)], \end{aligned} \tag{16.b}$$

mit der man dann für $\mathfrak{S}(X, Y)$ die Darstellung

$$\mathfrak{S}(X, Y) = \mathfrak{S}(X)\mathfrak{S}(Y) + \varDelta[X, Y] \tag{16.c}$$

erhält.

§ 111. Um nun noch die $\varPhi$-Reihen einzuführen, setzen wir mit den Parametern c, h, e, k die Hilfsargumente

$$u = h(x - c), \qquad U = h(X - c), \tag{17.a}$$

$$v = k(y - e), \qquad V = k(Y - e), \tag{17.b}$$

sowie die Reihen

$$\operatorname{sg}(X - x) = \sum_p \mathfrak{R}(u)_p \varPhi(U)_p,$$

$$\operatorname{sg}(Y - y) = \sum_q \mathfrak{R}(v)_q \varPhi(V)_q$$

an, in denen die Summation nach p und q von Null an zu laufen hat. Daraus ergeben sich dann mit den Abkürzungen

$$D(c, h)_p = \mathfrak{D}[\mathfrak{R}(u)_p], \qquad D(e, k)_q = \mathfrak{D}[\mathfrak{R}(v)_q], \tag{18}$$

$$D(c, h, e, k)_{pq} = \mathfrak{D}[\mathfrak{R}(u)_p \mathfrak{R}(v)_q], \tag{19}$$

$$\varDelta(p, q) = D(c, h, e, k)_{pq} - D(c, h)_p D(e, k)_q \tag{20}$$

die Reihen

$$2\mathfrak{S}(X) - 1 = \sum_p D(c, h)_p \varPhi(U)_p, \tag{21.a}$$

$$2\mathfrak{S}(Y) - 1 = \sum_q D(e, k)_q \varPhi(V)_q, \tag{21.b}$$

$$4\varDelta[X, Y] = \sum_p \sum_q \varDelta(p, q) \varPhi(U)_p \varPhi(V)_q. \tag{22}$$

Bildet man endlich noch von $E(X, x)$ den Durchschnitt nach der Verteilung von $K(x)_y$, so erhält man zunächst den Ausdruck

$$\mathfrak{D}_x[E(X, x)] = \sum_x E(X, x)\, U(x)_y.$$

Die rechte Seite liefert aber die zu $K(x)_y$ gehörige Summenfunktion $\mathfrak{S}(X)_y$, also ist

$$\mathfrak{S}(X)_y = \mathfrak{D}_x[E(X, x)], \tag{23.a}$$

$$2\mathfrak{S}(X)_y - 1 = \mathfrak{D}_x[\operatorname{sg}(X - x)], \tag{23.b}$$

woraus mit der Abkürzung

$$D(c, h, y)_p = \mathfrak{D}_x[\mathfrak{R}(u)_p] \tag{24}$$

die zu den Parametern c, h gehörende Darstellung

$$2\mathfrak{S}(X)_y - 1 = \sum_p D(c, h, y)_p \varPhi(U)_p \tag{25}$$

folgt.

Hält man die Gleichungen (13.a), (9) und (23.a) zusammen, so kann man ansetzen

$$\mathfrak{S}(X) = \mathfrak{D}[E(X, x)] = \mathfrak{D}_y[\mathfrak{D}_x[E(X, x)]] = \mathfrak{D}_y[\mathfrak{S}(X)_y]$$

oder

$$\mathfrak{S}(X) = \underset{y}{\mathfrak{D}}[\mathfrak{S}(X)_y]. \tag{26}$$

Wiederholt man die vorstehende Rechnung, indem man jedoch statt $E(X, x)$ das Polynom $\mathfrak{R}(u)_p$ zugrunde legt, so erhält man nach den eingeführten Bezeichnungen

$$D(c, h)_p = \underset{y}{\mathfrak{D}}[D(c, h, y)_p]. \tag{27}$$

Beachtet man ferner, daß — je nach der Art der Verteilung — aus der Summenfunktion durch Differenzenbildung oder durch Differentiation die Verteilungsfunktionen $\mathfrak{U}$ oder $\mathfrak{V}$ entspringen, so erhält man aus (26) zwischen den Verteilungsfunktionen von $K(x)$ und $K(x)_y$ für den Fall der Unstetigkeit die Beziehung

$$\mathfrak{U}(X) = \underset{y}{\mathfrak{D}}[\mathfrak{U}(X)_y], \tag{28.a}$$

und für den Fall der Stetigkeit

$$\mathfrak{V}(X) = \underset{y}{\mathfrak{D}}[\mathfrak{V}(X)_y]. \tag{28.b}$$

Hiernach kommt der Mischungsprozeß, der von den Stücken

$$S(X)_y, \quad \mathfrak{U}(x)_y, \quad \mathfrak{V}(x)_y, \quad D(c, h, y)_p$$

zu den Stücken

$$\mathfrak{S}(x), \quad \mathfrak{U}(x), \quad \mathfrak{V}(x), \quad D(c, h)_p$$

führt, darauf hinaus, daß man aus den zu $K(x)_y$ gehörenden Stücken mit den Gewichten $U(x)_y$ das Mittel nimmt.

§ 112. Bedeuten jetzt c und s den Argumentdurchschnitt und die Streuung der Mischung $K(x)$, so erhält man nach (11), weil die Größen x und $x - c$ nicht von y abhängen, die Gleichungen

$$\mathfrak{D}(x) = c, \tag{29}$$

$$\mathfrak{D}[(x - c)^2] = s^2. \tag{30}$$

Bedeuten ferner c_y und s_y die Argumentdurchschnitte und die Streuungen für die Teilreihen $K(x)_y$, so ist

$$\underset{x}{\mathfrak{D}}(x) = c_y, \tag{31}$$

$$\underset{x}{\mathfrak{D}}[(x - c_y)^2] = s_y^2. \tag{32}$$

Aus (29) und (31) folgt

$$c = \mathfrak{D}(x) = \underset{y}{\mathfrak{D}}[\underset{x}{\mathfrak{D}}(x)] = \underset{y}{\mathfrak{D}}(c_y). \tag{33}$$

Man kann nun die c_y als die Glieder einer Kollektivreihe ansehen, deren Argument die Werte der c_y selber sind, wobei die rH. eines

jeden Gliedes c_y durch $U(y)$, also durch die Verteilung von $K(y)$ bestimmt sein soll. Dann besagt die aus (33) folgende Beziehung

$$c = \underset{y}{\mathfrak{D}}(c_y), \tag{34}$$

daß c der Argumentdurchschnitt zu der Reihe der c_y ist. Zu dieser Reihe gehört ferner eine Streuung $\operatorname{str}(c_y)$, die durch die Gleichung

$$\operatorname{str}(c_y)^2 = \underset{y}{\mathfrak{D}}[(c_y - c)^2] \tag{35}$$

gegeben ist. Setzt man nun die Identität

$$(x-c)^2 = (x-c_y)^2 + 2(x-c_y)(c_y-c) + (c_y-c)^2$$

an und führt die zu $K(x, y)$ gehörige $\mathfrak{D}$-Operation aus, so erhält man links wegen (30) den Wert s^2. Rechts erhält man dagegen unter Berücksichtigung von (32) und (10) die drei Bestandteile

$$\mathfrak{D}[(x-c_y)^2] = \underset{y}{\mathfrak{D}}[\underset{x}{D}[(x-c_y)^2]] = \underset{y}{\mathfrak{D}}(s_y^2),$$

$$2\mathfrak{D}[(x-c_y)(c_y-c)] = 2\underset{y}{\mathfrak{D}}[(c_y-c)\underset{x}{\mathfrak{D}}(x-c_y)],$$

$$\mathfrak{D}[(c_y-c)^2] = \underset{y}{\mathfrak{D}}[(c_y-c)^2] = \operatorname{str}(c_y)^2.$$

Der zweite Bestandteil verschwindet wegen (31), folglich wird

$$s^2 = \underset{y}{\mathfrak{D}}(s_y^2) + \operatorname{str}(c_y)^2. \tag{36}$$

Die Streuung von $K(x)$ setzt sich also aus zwei Termen zusammen, von denen der eine durch die Streuungen der Reihen $K(x)_y$, der andere dagegen durch die Argumentdurchschnitte derselben Reihen erzeugt wird. Wie sich diese Art der Zusammensetzung jedesmal äußert, hängt natürlich von der Gruppierung der Größen c_y und s_y ab; sie kann unter Umständen bewirken, daß die charakteristischen Merkmale der $K(x)_y$ durch den Mischungsprozeß vollständig verwischt werden. Sind z. B. die s_y durchweg klein, so besitzen die Verteilungskurven der $K(x)_y$ steile Gipfel; trotzdem wird die Verteilungskurve von $K(x)$ stark abgeflacht verlaufen, sobald die Werte der c_y stark auseinandergehen, weil dann der große Betrag von $\operatorname{str}(c_y)$ in (36) auch der Streuung s einen großen Wert verleiht.

Führt man jetzt noch die Parameterwerte h und h_y durch die Gleichungen

$$1 = 2(hs)^2, \qquad 1 = 2(h_y s_y)^2$$

ein, so liefern die in (18) angesetzten Ausdrücke

$$D(c, h)_p = \mathfrak{D}[\mathfrak{R}(u)_p] \tag{37}$$

zusammen mit den Größen c und s die numerischen Elemente der Mischung $K(x)$, wobei dann nach (27)

$$D(c, h)_p = \underset{y}{\mathfrak{D}}[c, h, y)_p] \tag{38}$$

wird. Drückt man hierin nun noch die rechter Hand auftretenden und zu $K(x)_y$ gehörenden Koeffizienten $D(c, h, y)_p$ nach den früher entwickelten Transformationsgleichungen durch ihre normalen Werte $D(c_y, h_y, y)_p$ aus, so ist damit der direkte Zusammenhang zwischen den numerischen Elementen der Mischung $K(x)$ und den Elementen der Mischungsbestandteile $K(x)_y$ hergestellt.

Da die in der angegebenen Weise aus (38) folgenden Gleichungen keine durch Einfachheit bemerkenswerten Resultate liefern, so soll der allgemeine Fall hier nicht weiter verfolgt werden. Dagegen ist es von Interesse, noch einen besonderen Fall zu betrachten.

§ 113. Bei der Einführung des hier untersuchten Mischungsprozesses war erwähnt worden, daß dem Prozeß im allgemeinen die Tendenz innewohnt, von dem E.-G. hinwegzuführen. Um die Berechtigung dieser Bemerkung deutlich zu machen, genügt der Nachweis, daß das E.-G., wenn es für die einzelnen Teilreihen $K(x)_y$ gilt, durch die Mischung in $K(x)$ im allgemeinen zerstört wird, daß also die Erhaltung des E.-G. nur unter besonderen Umständen stattfindet. Indem wir die Bedeutung, die den Buchstaben c, s, h und c_y, s_y, h_y im letzten Paragraphen beigelegt wurde, festhalten, wollen wir annehmen, daß die Reihen $K(x)_y$ sämtlich dem E.-G. folgen. Setzt man dann den Ausdruck

$$\begin{aligned} T &= \exp\left[-2h_y(x-c_y)s_y v - (s_y v)^2\right] \\ &= \sum_p \mathfrak{R}(h_y x - h_y c_y)_p (2 s_y v)^p \end{aligned}$$

an, so liefert der nach der Verteilung von $K(x)_y$ genommene Durchschnitt die Gleichung

$$\underset{x}{\mathfrak{D}}(T) = \sum_p \underset{x}{\mathfrak{D}}[\mathfrak{R}(h_y x - h_y c_y)_p](2 s_y v)^p. \tag{39}$$

Die $\mathfrak{D}$-Größen rechter Hand sind aber nichts anderes, als die $\mathfrak{D}$-Koeffizienten in der zu $K(x)_y$ gehörigen Φ-Reihe, und zwar für die Normalform dieser Reihe. Diese Koeffizienten müssen nun mit Ausnahme des ersten sämtlich verschwinden, wenn die $K(x)_y$ sämtlich dem E.-G. gehorchen sollen. Die Reihe (39) reduziert sich also auf das zu $p = 0$ gehörige Anfangsglied, woraus

$$\underset{x}{\mathfrak{D}}(T) = 1 \tag{40}$$

folgt.

Weiter können wir wegen $hs = h_y s_y$ schreiben

$$T = \exp\left[-2h(x - c_y)sv - (s_y v)^2\right],$$

woraus sich mit der Abkürzung

$$E = \exp\left[- 2h(c_y - c)sv + (s_y^2 - s^2)v^2\right]$$

die Gleichung

$$\exp\left[- 2h(x - c)sv - s^2v^2\right] = TE$$

und, wenn man links nach v entwickelt,

$$\sum_p \mathfrak{R}(hx - hc)_p (2sv)^p = TE$$

ergibt. Nimmt man hiervon den Durchschnitt nach der Verteilung von $K(x, y)$, so erhält man wegen (37) auf der linken Seite die Reihe

$$\sum_p D(c, h)_p (2sv)^p,$$

worin die $D(c, h)_p$ die D-Koeffizienten zu der Normalform von $K(x)$ sind. Auf der rechten Seite entsteht, da der Faktor E von x frei ist, unter Berücksichtigung von (40) der Reihe nach

$$\mathfrak{D}(TE) = \mathfrak{D}_y[\mathfrak{D}_x(TE)] = \mathfrak{D}_y[E\mathfrak{D}_x(T)] = \mathfrak{D}_y(E).$$

Demnach ist wegen $hs = 1 : \sqrt{2}$

$$\sum_p D(c, h)_p (2sv)^p = \mathfrak{D}_y[\exp(-\sqrt{2}(c_y - c)v + (s_y^2 - s^2)v^2)]. \tag{41}$$

Entwickelt man rechts nach Potenzen von v, so verschwinden die Koeffizienten nicht identisch, abgesehen von den beiden Gliedern mit v und v^2. Folglich können die Koeffizienten $D(c, h)_3$, $D(c, h)_4$, ... nicht allgemein verschwinden, wie es doch nötig wäre, wenn auch für $K(x)$ stets das E.-G. gelten soll; vielmehr ist dazu nötig, daß die numerischen Elemente der $K(x)_y$, d. h. hier die Größen c_y und s_y bestimmten Bedingungen genügen, die — zunächst wenigstens — in unendlicher Anzahl auftreten. Das E.-G. der Reihen $K(x)_y$ wird also in der Mischung $K(x)$ im allgemeinen zerstört.

§ 114. Der betrachtete Fall soll jetzt noch weiter verfolgt werden, und zwar unter der Voraussetzung, daß die c_y miteinander und dann auch natürlich mit c zusammenfallen. Diese Voraussetzung ist deswegen von Interesse, weil sie — mit größerer oder geringerer Annäherung — bei Beobachtungsfehlern vorzukommen pflegt.

Für $c_y = c$ entsteht aus (41)

$$\sum_p D(c, h)_p (2sv)^p = \mathfrak{D}_y[\exp(s_y^2v^2 - s^2v^2)], \tag{42}$$

wozu noch unter Berücksichtigung von (36) die Gleichung

$$s^2 = \mathfrak{D}_y(s_y^2) \tag{43}$$

hinzukommt. Entwickelt man (42) rechts nach v und vergleicht mit

der linken Seite, so findet man zunächst für $D(c, h)_2$ den Wert Null, wie vorherzusehen war. Ferner wird

$$32 s^4 D(c, h)_4 = \underset{y}{\mathfrak{D}}[(s_y^2 - s^2)^2], \tag{44.a}$$

$$384 s^6 D(c, h)_6 = \underset{y}{\mathfrak{D}}[(s_y^2 - s^2)^3], \tag{44.b}$$

usw.,

während die ungeraden $D(c, h)$ durchweg verschwinden. Damit nimmt die Gleichung für die Summenfunktion der Mischung $K(x)$ unter Benutzung des normalen Hilfsarguments

$$u = h(x - c)$$

die Gestalt

$$2\mathfrak{S}(x) - 1 = \Phi(u) + D(c, h)_4 \Phi(u)_4 + D(c, h)_6 \Phi(u)_6 + \cdots \tag{45.a}$$

an, woraus für die zugehörige Verteilungsfunktion $\mathfrak{V}(x)$ die Darstellung

$$2\mathfrak{V}(x) = h\Phi(u)_1 + D(c, h)_4 h\Phi(u)_5 + \cdots \tag{45.b}$$

folgt. Das erste Glied rechts entspricht einem E.-G., dessen Verteilungsfunktion $\mathfrak{E}(x)$ durch die Gleichung

$$2\mathfrak{E}(x) = h\Phi(u)_1 \tag{46}$$

gegeben ist und mit der zu $K(x)$ gehörigen Funktion $\mathfrak{V}(x)$ den Argumentdurchschnitt und die Streuung gemeinsam hat. Betrachtet man nun $\mathfrak{E}(x)$ als eine Art idealer Norm, auf die $\mathfrak{V}(x)$ zu beziehen ist, so kann man, je nachdem $\mathfrak{V} - \mathfrak{E}$ positiv, null oder negativ ist, die durch $\mathfrak{V}(x)$ bestimmten Häufigkeiten von x als *übernormal*, *normal* oder *unternormal* bezeichnen. Die hierin liegende Charakterisierung von $\mathfrak{V}(x)$ hängt offenbar von dem Vorzeichen der Größe

$$2(\mathfrak{V} - \mathfrak{E}) = D(c, h)_4 h\Phi(u)_5 + D(c, h)_6 h\Phi(u)_7 + \cdots \tag{47}$$

ab. Beschränkt man sich für den Augenblick auf das Glied mit Φ_5, dessen Koeffizient nach (44.a) wesentlich positiv ist, und beachtet man ferner den aus den Φ-Tafeln folgenden Gang von Φ_5, so erkennt man, daß $\mathfrak{V}$ übernormale Werte annimmt, wenn der absolute Betrag von u hinreichend klein oder hinreichend groß genommen wird, daß dagegen für mittelgroße u unternormale Werte auftreten.

Die vorstehende Bemerkung über das Verhalten von $\mathfrak{V}(x)$ stützt sich nur auf das erste Reihenglied in (47), könnte also durch die Heranziehung der übrigen Glieder wieder umgestoßen werden. Es soll nun gezeigt werden, daß dies nicht eintritt.

§ 115. Denkt man sich der Einfachheit halber den Nullpunkt so verschoben, daß die c-Größen sämtlich verschwinden, so kommt der Voraussetzung nach jeder Reihe $K(x)_y$ ein Verteilungsgesetz $\mathfrak{V}(x)_y$ zu, das durch den Ausdruck

$$\mathfrak{V}(x)_y = h_y \exp(-h_y^2 x^2) : \sqrt{\pi}$$

gegeben ist. Daraus entspringt für $\mathfrak{B}(x)$ die Darstellung

$$\mathfrak{B}(x) = \underset{y}{\mathfrak{D}}[\mathfrak{B}(x)_y] = \sum_y U(y)\mathfrak{B}(x)_y$$

oder

$$\sqrt{\pi}\,\mathfrak{B}(x) = \sum_y U(y) h_y \exp(-h_y^2 x^2). \tag{48}$$

Da nun für den Charakter von $\mathfrak{B}(x)$ das Vorzeichen der Größe

$$F(x) = \sqrt{\pi}\,[\mathfrak{B}(x) - \mathfrak{E}(x)]$$

entscheidend ist, so handelt es sich um das Verhalten des Ausdrucks

$$F(x) = -h \exp(-h^2 x^2) + \sum_y U(y) h_y \exp(-h_y^2 x^2). \tag{49}$$

Schreibt man dafür

$$F(x) \exp(-h^2 x^2) = -h + \sum_y U(y) h_y \exp[(h^2 - h_y^2) x^2],$$

so kommen in der Summe rechts positive Glieder vor, die mit wachsendem x über alle Grenzen gehen, da ja nach der aus (43) folgenden Gleichung

$$h^{-2} = \underset{y}{\mathfrak{D}}(h_y^{-2}) = \sum_y U(y) h_y^{-2} \tag{50}$$

der Parameter h zwischen dem größten und kleinsten h_y liegen muß. Demnach ist für hinreichend große x der Ausdruck F sicher positiv, also $\mathfrak{B}(x)$ sicher übernormal.

Setzt man in (49) $x = 0$, so wird

$$F(0) = -h + \sum_y U(y) h_y. \tag{51}$$

Setzt man ferner

$$h : h_y = z_y, \qquad F(0) = hG, \tag{52}$$

so wird

$$G = -1 + \sum_y U(y) z_y^{-1}, \tag{53}$$

und wegen (50)

$$\sum_y U(y) z_y^2 - 1 = 0. \tag{54}$$

Es soll nun gezeigt werden, daß G notwendig positiv ist.

Zu dem Ende betrachten wir für den Augenblick G als eine Funktion der Veränderlichen z_y, wobei wegen der Gleichung $h = h_y z_y$ negative Werte der z_y ausgeschlossen bleiben sollen, und suchen unter Berücksichtigung der Bedingung (54) das Minimum von G auf.

§ 116. Der Einfachheit halber soll zunächst vorausgesetzt werden, daß die Reihe $K(y)$ unstetig sei, daß also für y in Wirklichkeit nur eine endliche Anzahl diskreter Argumentwerte mit der Verteilungsfunktion $\mathfrak{U}(y)$ in Betracht komme. Dann sind die Gleichungen (53) und (54) in der Gestalt

$$G = -1 + \sum_y \mathfrak{U}(y) z_y^{-1}, \tag{55}$$

$$\sum_y \mathfrak{U}(y) z_y^2 - 1 = 0 \tag{56}$$

zu schreiben, wo die Summen nur über die in $K(y)$ auftretenden Argumente auszudehnen sind.

Das Gebiet der zulässigen z umfaßt, da die $\mathfrak{U}$-Größen wesentlich positiv sind, wegen (56) nur endliche z. Da ferner negative z ausgeschlossen sind, so kann G nicht negativ unendlich werden, weil dazu eines der z negativ unendlich klein werden müßte. Es existiert also für G eine endliche untere Grenze. Da überdies für unendlich kleine positive z der Ausdruck G positiv unendlich wird, so gehören die endlichen Werte von G solchen Stellen im Gebiete der z an, in denen G als Funktion der z regulär verläuft. Folglich besitzt G sicher ein Minimum, das nach den gewöhnlichen Regeln aufgefunden werden kann. Demgemäß setzen wir mit einem Multiplikator L den Ausdruck

$$G + L(\sum_y \mathfrak{U}(y) z_y^2 - 1)$$

an, nehmen davon die partiellen Ableitungen erster Ordnung nach den z_y und setzen diese Ableitungen einzeln gleich Null. Daraus entsteht das Gleichungssystem

$$0 = -\mathfrak{U}(y) z_y^{-2} + 2 L \mathfrak{U}(y) z_y$$

oder

$$1 = 2 L z_y^3,$$

aus dem zu schließen ist, daß an dem Orte des Minimums die z_y einen und denselben Wert besitzen, der mit z bezeichnet werden möge. Ersetzt man jetzt in (55) und (56) die z_y durch z und beachtet, daß die Summe der $\mathfrak{U}(y)$ gleich Eins ist, so wird

$$G = -1 + z^{-1}, \qquad z^2 - 1 = 0.$$

Aus der zweiten Gleichung folgt $z = 1$, und damit aus der ersten als Minimum von G der Wert $G = 0$.

Das Zusammenfallen der z_y zieht nach (52) das der h_y nach sich, d. h. der Wert $G = 0$ kann überhaupt nur auftreten, wenn die E.-G., die den einzelnen Reihen $K(x)_y$ zukommen, vollständig übereinstimmen. Dieser Fall kommt nun aber hier garnicht in Betracht, denn die ganze Rechnung hatte ja überhaupt nur dann einen Zweck, wenn es sich um die Mischung von ungleichen Verteilungen handelte. Werden demnach die h_y als ungleich vorausgesetzt, so kann der Wert von G nicht zu dem Minimum herabsinken, ist also positiv, und folgeweise auch der Wert von $F(0)$. Daraus ergibt sich weiter, daß $\mathfrak{B}(x)$ für $x = 0$ und — der Stetigkeit wegen — auch für hinreichend kleine Werte von x übernormal verläuft.

Die zwischen den Grenzen $\pm \infty$ genommenen Integrale über $\mathfrak{B}(x)$ und $\mathfrak{E}(x)$ sind gleich Eins, woraus für das entsprechende Integral über die Differenz $\mathfrak{B} - \mathfrak{E}$ der Wert Null folgt. Danach müssen in dem Integral über $\mathfrak{B} - \mathfrak{E}$ die positiven Beträge dieser Differenz, die

für gewisse Werte von x auftreten, darch negative Beträge für andere x kompensiert werden.

Das vorstehende, zunächst für ein unstetiges $K(y)$ gefundene Ergebnis darf nun auf den Fall eines stetigen $K(y)$ übertragen werden, da sich die beobachteten stetigen Kollektivreihen für die numerische Rechnung als unstetige Reihen behandeln lassen. Damit gelangen wir zu folgendem Satz: Wird aus einfachen Exponentialgesetzen mit zusammenfallenden Argumentdurchschnitten das gemischte Gesetz $\mathfrak{B}(x)$ gebildet, so ist $\mathfrak{B}(x)$ in der Umgebung des Durchschnittes $\mathfrak{D}(x)$ und ebenso in hinreichend großer Entfernung davon sicher übernormal, während in mittelgroßer Entfernung von $\mathfrak{D}(x)$ sicher unternormale Werte auftreten.

§ 117. Die Wirkungen des hier betrachteten Mischungsprozesses legen den Gedanken nahe, ob man nicht ein beliebig vorgelegtes Verteilungsgesetz aus einfacheren Formen, z. B. aus bloßen Exponentialgesetzen „zusammenmischen" könne. Zur Verdeutlichung der gestellten Frage möge folgendes Beispiel dienen. Wenn für einen gegebenen K.-G. die Darstellung

$$u = h(x - c), \quad 2\mathfrak{S}(x) - 1 = \sum_p D(c, h)_p \Phi(u)_p \tag{57}$$

vorliegt, so denke man sich zunächst den Ausdruck

$$m_1 \Phi[h_1(x - c_1)] + m_2 \Phi[h_2(x - c_2)] \tag{58}$$

angesetzt und letzteren unter Benutzung der in § 98—99 entwickelten Formeln nach dem Hilfsargument u in die Reihenform

$$\sum_p A_p \Phi(u)_p \tag{59}$$

gebracht. Dann hängen die Koeffizienten A_p von den sechs Größen

$$m_1, \quad m_2, \quad c_1, \quad c_2, \quad h_1, \quad h_2$$

ab, und man kann sich vornehmen, diese Größen so zu wählen, daß die Reihe (59) mit der Reihe (57) in den sechs ersten Gliedern übereinstimmt. In ähnlicher Weise kann man, wenn statt (58) ein aus drei Gliedern gebildeter Ausdruck mit neun verfügbaren Parametern zugrunde gelegt wird, Übereinstimmung zwischen (57) und (59) in den neun ersten Gliedern herbeizuführen suchen usw., bis alle in (57) merklichen Glieder erschöpft sind.

Es liegt auf der Hand, daß Ausdrücke von der Form (58) oder von einer nötigenfalls erweiterten Gestalt, falls sie sich glatt herstellen ließen, für die Rechnung manche Bequemlichkeiten bieten würden. In Wirklichkeit zeigt sich indessen, daß die Auflösung der Gleichungen

$$A_p = D(c, h)_p$$

schon bei nur sechs Unbekannten recht unbequem wird, und daß imaginäre Lösungen nicht ausgeschlossen sind. Wesentlicher als dieser Umstand ist jedoch der andere, daß die gedachte Zerlegung meistens nur eine formale Bedeutung und gar keine innere Beziehung zu dem darzustellenden K.-G. besitzt. Man wird daher solche „Entmischungen" nur da versuchen, wo die besonderen Eigenschaften eines K.-G. bestimmte Fingerzeige dafür an die Hand geben.

§ 118. Die Sätze, die wir bei der allgemeinen Untersuchung der Kollektivreihe $K(x, y)$ gewonnen haben, schließen auch, wie sich zeigen wird, die Beantwortung der Frage nach der sogenannten Unabhängigkeit der Argumente in sich. Deshalb möge dieser Gegenstand hier sogleich mit behandelt werden.

Zu dem Ende gehen wir wieder auf die Kollektivreihe $K(x, y)$ und die daraus gebildeten Teilreihen $K(x)_y$ nebst den dazu gehörigen Verteilungsgrößen $U(x, y)$ und $U(x)_y$ zurück und halten uns dabei gegenwärtig, daß der Verlauf der Summenkurve, die zu jedem $K(x)_y$ gehört, durch den Verlauf der entsprechenden Größen $U(x)_y$ vorgeschrieben ist. Bezeichnet man nun für den Augenblick die zum Index y oder zur Teilreihe $K(x)_y$ gehörige Summenkurve mit $C(y)$, so sind zunächst folgende zwei Fälle ins Auge zu fassen:

a) die $C(y)$ stimmen nach Form und Lage im wesentlichen überein, d. h. die zwischen ihnen etwa vorhandenen Unterschiede sind so klein, daß man sie unbedenklich als den Rest unausgeglichener Zufälligkeiten ansehen kann;

b) die $C(y)$ weisen ausgesprochene Unterschiede auf, die einen deutlich von dem Index y abhängenden Gang besitzen.

Im ersten Falle sagt man, daß das Argument x *unabhängig* von dem Argument y variiere, wogegen im zweiten Falle x als von y *abhängend* bezeichnet wird. Darüber hinaus kann nun aber auch noch der Fall eintreten, daß sich eine bündige Entscheidung weder für a) noch für b) geben läßt; man hat dann die Frage, ob für x Abhängigkeit oder Unabhängigkeit vorhanden sei, vorläufig offen zu lassen.

Von den angeführten drei Möglichkeiten betrachten wir, unter Ausscheidung der dritten, zunächst die erste, um nachher noch die zweite kurz zu berühren.

§ 119. Soll Unabhängigkeit vorhanden sein, so müssen die $U(x)_y$, aus denen die Summenkurven der Teilreihen $K(x)_y$ entspringen, für jeden Indexwert y auf dieselbe Funktion von x führen, müssen also gleich einer gewissen, nur von x abhängenden Funktion $P(x)$ sein. Damit nimmt die früher aufgestellte Gleichung

$$U(x, y) = U(x)_y \, U(y) \tag{60}$$

die Gestalt

$$U(x, y) = P(x) \, U(y) \tag{61}$$

an. Summiert man die letzte Gleichung nach y, so liefert links die Summe über die $U(x, y)$ den Ausdruck $U(x)$, während rechter Hand die Summe über die $U(y)$ den Wert Eins annimmt. Demnach wird

$$U(x) = P(x),$$

womit dann (61) in

$$U(x, y) = U(x)\, U(y) \tag{62}$$

übergeht.

Die vorstehende Bedingung, die sich als eine notwendige Folge der vorausgesetzten Unabhängigkeit erwiesen hat, ist umgekehrt auch hinreichend. Denn wenn in einer vorgelegten Reihe $K(x, y)$ die Verteilungsgrößen $U(x, y)$ in der Gestalt (62) angesetzt werden können, so erhält man durch Vergleichung mit (60) für die $U(x)_y$ den vom Index y freien Ausdruck $U(x)$, woraus dann weiter die Übereinstimmung aller Summenkurven $C(y)$ folgt.

Aus dem Umstande, daß die Bedingung (62) eine nach x und y symmetrische Gestalt besitzt, fließt noch die Bemerkung, daß wenn x von y unabhängig ist, umgekehrt auch y nicht von x abhängt.

Bedeuten nun weiter $S(x)$ und $T(y)$ irgend zwei Ausdrücke, die jedesmal nur von dem einen hingeschriebenen Argument x oder y abhängen, so wird unter Berücksichtigung von (62)

$$\mathfrak{D}(ST) = \sum_x \sum_y S\,T\,U(x, y) = [\sum_x S\,U(x)] \cdot [\sum_y T\,U(y)].$$

Andrerseits darf man in diesem Falle nach (10) und (11) ansetzen

$$\mathfrak{D}(S) = \sum_x S\,U(x), \quad \mathfrak{D}(T) = \sum_y T\,U(y),$$

so daß

$$\mathfrak{D}(ST) = \mathfrak{D}(S) \cdot \mathfrak{D}(T)$$

wird. Diese Gleichung wollen wir noch etwas anders fassen. Führt man den Ausdruck

$$W = \mathfrak{D}(ST) - \mathfrak{D}(S) \cdot \mathfrak{D}(T) \tag{63}$$

ein, so fordert die Unabhängigkeit, daß W durchweg null sei. Nun wird aber bei einer beobachteten Reihe $K(x, y)$, wenn man die Werte von W für die verschiedenen Wertepaare von x und y berechnet, selbst im Falle wirklicher Unabhängigkeit kein beständiges Verschwinden von W stattfinden, weil ja in den $\mathfrak{D}$-Größen des Ausdruckes W Reste von unausgeglichenen Zufälligkeiten stehen bleiben. Es kommt also nicht darauf an, daß die Werte von W wirklich verschwinden, sondern nur darauf, daß diese Werte, die wir jetzt als die zwischen dem Kriterium und der Beobachtung übrigbleibenden Widersprüche bezeichnen wollen, hinreichend klein sind. Demnach ist das Kriterium als erfüllt anzusehen, wenn die Widersprüche nur so groß sind, daß sie unausgeglichenen Zufälligkeiten zur Last gelegt werden dürfen.

§ 120. Setzt man in (63) für S und T andre und andre Ausdrücke ein, so nimmt das für die Unabhängigkeit aufgestellte Kriterium andre und andre Gestalten an, zwischen denen man je nach den Umständen wählen wird.

Setzt man wie früher

$$2E(X, x) = \operatorname{sg}(X - x) + 1,$$
$$\varDelta[X, Y] = \mathfrak{S}(X, Y) - \mathfrak{S}(X) \cdot \mathfrak{S}(Y),$$

so führt die Substitution $S = E(X, x)$, $T = E(Y, y)$ zu der Form

$$W = \varDelta[X, Y]. \tag{64}$$

Diese Form ist sehr bequem, weil man die $\mathfrak{S}$-Größen unmittelbar aus der Verteilungstafel des vorgelegten K.-G. ableiten und damit die Widersprüche für passend gewählte Wertepaare (XY) herstellen kann. Zugleich vermag man wegen der Bedeutung von $\varDelta[X, Y]$ sofort zu überblicken, ob Werte wie z. B. $W = 0.01$ oder $W = 0.001$ als erheblich anzusehen sind oder nicht.

Ist der vorgelegte K.-G durchweg stetig, so entsteht aus (64) durch Differentiieren nach x und y die Form

$$W = \mathfrak{V}(X, Y) - \mathfrak{V}(X) \cdot \mathfrak{V}(Y), \tag{65}$$

welche die Kenntnis der zu $K(x, y)$, $K(x)$ und $K(y)$ gehörigen stetigen Verteilungsfunktionen $\mathfrak{V}(x, y)$, $\mathfrak{V}(x)$ und $\mathfrak{V}(y)$ voraussetzt, also für gewöhnlich nicht in Betracht kommen wird, da die $\mathfrak{V}$ bei beobachteten Reihen immer erst aus der interpolatorischen Darstellung gefunden werden.

Ist $K(x, y)$ in beiden Argumenten unstetig, so folgt aus (64) durch Differenzenbildung oder auch direkt aus (62) die Form

$$W = \mathfrak{U}(X, Y) - \mathfrak{U}(X) \cdot \mathfrak{U}(Y), \tag{66}$$

deren Bestandteile sich unmittelbar aus der Verteilungstafel bilden lassen. Trotzdem ist im allgemeinen die Form (64) vorzuziehen, denn wenn die Menge der für x und y beobachteten Wertsysteme einigermaßen groß ist, so besitzen die $\mathfrak{U}$-Größen vorwiegend kleine Werte, und man vermag meistens nicht ohne weiteres zu übersehen, ob die Werte der Widersprüche (66) als erheblich zu betrachten sind oder nicht.

Setzt man wie früher die Ausdrücke

$$u = h(x - c), \quad v = k(y - e),$$
$$D(c, h)_p = \mathfrak{D}[R(u)_p], \quad D(e, k)_q = \mathfrak{D}[R(v)_q],$$
$$D(c, h, e, k)_{pq} = \mathfrak{D}[R(u)_p R(v)_q],$$
$$\varDelta(p, q) = D(c, h, e, k)_{pq} - D(c, h)_p D(e, k)_q$$

an, so erhält man mit $S = R(u)_p$, $T = R(v)_q$ die Form

$$W = \Delta(p, q). \tag{67}$$

Für $p = 0$ oder $q = 0$ verschwinden die $\Delta(p, q)$, wie leicht zu sehen, identisch, so daß die vorstehenden W-Größen nur für

$$p = 1, 2, \ldots \text{ und } q = 1, 2, \ldots$$

anzusetzen sind. Da die $\Delta(p, q)$ in (22) als die Koeffizienten der Φ-Reihe für $\Delta[X, Y]$ auftreten, so kann man an der Hand der Φ-Tafeln sofort übersehen, wie die Werte der $\Delta(p, q)$ auf die Verteilung wirken. Indessen empfiehlt sich wegen der mit der Berechnung der $\Delta(p, q)$ verbundenen Arbeit die Benutzung von (67) nur dann, wenn man jene Größen noch aus anderen Gründen zu ermitteln Anlaß hat.

Setzt man endlich $S = x^p$, $T = y^q$, so wird

$$W = \mathfrak{D}(x^p y^q) - \mathfrak{D}(x^p) \cdot \mathfrak{D}(y^q). \tag{68}$$

Diese Form von W ist in analytischer Hinsicht die einfachste und darum für gewisse analytische Untersuchungen als Ausgangspunkt ganz gut geeignet. Sie gestattet aber — abgesehen etwa von dem Falle $p = q = 1$ — kein unmittelbares Urteil darüber, ob die durch (68) gegebenen Widersprüche als erheblich anzusehen sind oder nicht.

§ 121. Sind die Widersprüche, die das benutzte Kriterium liefert, unerheblich, so wird man die Unabhängigkeit als nachgewiesen ansehen dürfen. Dagegen darf man, wenn die Widersprüche erheblich sind, daraus noch nicht schließen, daß damit nun auch wirklich die Abhängigkeit festgestellt sei, denn dazu gehört, wie oben angegeben wurde, daß die Unterschiede der Summenkurven einen von dem Index y abhängigen Gang zeigen.

Tritt dies nicht ein, so ist auch die Abhängigkeit nicht nachgewiesen, und man kann dann nur sagen, daß der vorgelegte K.-G. für die verlangte Entscheidung entweder nicht zureicht, wie das ja bei zu geringem Umfang von $K(x, y)$ vorkommen kann, oder aber für eine solche Untersuchung überhaupt untauglich ist. Wollte z. B. ein Anthropologe eine Untersuchung über die gleichzeitigen Abmessungen von Ober- und Unterarm in der Weise anstellen, daß er die zu messenden Individuen ganz planlos ohne Rücksicht auf Rasse, Alter, Geschlecht und etwa vorhandene Mißbildungen und Verkrüppelungen zusammensuchte, so würde sich niemand darüber wundern, daß ein solches Material sich für die vorliegende Frage als unbrauchbar erweist.

Bei der Untersuchung der Abhängigkeit wird der Gang der Bearbeitung sich danach zu richten haben, ob man bis zur Aufstellung der Φ-Reihe gehen oder aber vorher bei mehr summarischen Ergebnissen stehen bleiben will. Zu dem ersten Falle ist zu bemerken, daß die numerischen Elemente der gemischten Reihen $K(x)$ und $K(y)$ mit

den Größen $\varDelta(p, q)$ zusammengenommen das vollständige System der numerischen Elemente zu der Reihe $K(x, y)$ enthalten. Hierbei zerfallen die Elemente von $K(x, y)$ in zwei Gruppen, deren erste die Elemente von $K(x)$ und $K(y)$ umfaßt, die für sich allein im Falle der Unabhängigkeit zur Charakteristik von $K(x, y)$ ausreichen, da ja dann die Größen $\varDelta(p, q)$ gleich Null zu setzen sind. Die zweite Gruppe dagegen umfaßt die Größen $\varDelta(p, q)$, die hier als die für die Abhängigkeit charakteristischen Bestimmungsstücke auftreten.[1]) Will man die Arbeit, die zur Berechnung der $\varDelta(p, q)$ erforderlich ist, nicht aufwenden, so empfiehlt es sich, wenigstens die Elemente von $K(x)$ und $K(y)$ in der durch die Umstände angezeigten Ausdehnung abzuleiten und darauf nach (64) das System der W-Größen graphisch oder tabellarisch in übersichtlicher Weise zum Vorschein zu bringen. Das Produkt $\mathfrak{S}(X)\mathfrak{S}(Y)$, das für sich allein ebenfalls als eine Summenfunktion aufgefaßt werden kann, dient dann als eine Art Norm, mit der die beobachtete Funktion $\mathfrak{S}(X, Y)$ verglichen wird.

Von dem Nachweise des Vorhandenseins einer Abhängigkeit ist die Frage nach ihrer Form wohl zu trennen. Wenn eine analytische Darstellung der Kollektivreihe $K(x, y)$ vorliegt, so ist allerdings die Form der Abhängigkeit dadurch gegeben, daß die numerischen Elemente der Teilreihe $K(x)_y$ als bekannte Funktionen des Index y erscheinen. Indessen ist darin nicht das enthalten, was man bei der Frage nach der Form der Abhängigkeit gewöhnlich wissen will. Wenn z. B. ein Kanon der menschlichen Gestalt nicht bloß nach dem Augenmaß, sondern auf Grund von planmäßig ausgeführten Messungen aufgestellt werden soll, so handelt es sich um folgende Aufgabe: aus Messungen ist eine Kollektivreihe gewonnen, mit der Körperlänge als Argument y und der Länge eines bestimmten Körpergliedes als Argument x; es soll gezeigt werden, daß dem Quotienten $x : y$ die Tendenz gegen einen bestimmten konstanten Wert innewohnt. Wenn sich nun zeigt, daß die Streuungen der Teilreihen $K(x)_y$ klein sind, und daß die Argumentdurchschnitte dieser Reihen merklich proportional

1) Bildet man aus den $\varDelta(p, q)$ einen Ausdruck Q, dessen Verschwinden das gleichzeitige Verschwinden aller $\varDelta(p, q)$ zur Folge hat, so könnte der Wert von Q nach der Bezeichnung von *Galton* und *Pearson* als sogenannter „Korrelationskoeffizient", d. h. als die durch eine einzige Zahl gegebene Maßbestimmung der Abhängigkeit dienen. Ein solches Q, für das man z. B. die Quadratsumme der $\varDelta$ wählen könnte, wäre indessen nur von untergeordnetem Werte, weil es über die *Form* der Abhängigkeit nichts aussagt. Ganz verfehlt wäre es natürlich, als schlechthin gültiges Maß der Abhängigkeit einen einzigen Korrelationskoeffizienten aufzustellen, der verschwinden kann, ohne daß die $\varDelta(p, q)$ gleichzeitig mit verschwinden. Vergleiche hierzu auch die Bemerkungen von *G. F. Lipps* in der Abhandlung „Die Bestimmung der Abhängigkeit zwischen den Merkmalen eines Gegenstandes" (Berichte der math.-phys. Klasse der k. Sächs. Ges. d. Wiss. 1905).

zu y laufen, so hat man damit allerdings das gefunden, was man sucht. In der Regel liegen jedoch die Umstände nicht so einfach, wie in dem angeführten Beispiele, denn die Funktion $f(x, y)$, deren Konstanz die Natur „anstrebt“, ist für gewöhnlich nicht im voraus bekannt. Ferner muß man, wenn der Umfang der Reihe nicht sehr groß ist, damit rechnen, daß Abhängigkeiten, die nicht sehr scharf ausgeprägt sind, durch die den einzelnen Gliedern der Reihe anhaftenden Zufälligkeiten mehr oder weniger verdeckt werden. Man kann sich darum auch nicht im voraus einen für alle Fälle passenden Algorithmus zurecht legen, sondern ist darauf angewiesen, die geeigneten Hilfsmittel von Fall zu Fall ausfindig zu machen.

Fünfzehnte Vorlesung.

Unsicherheit der numerischen Elemente.

§ 122. Nach der Untersuchung über die Mischung von Verteilungen nehmen wir die Betrachtung der Kollektivreihen, die nur ein einziges Argument x enthalten, wieder auf und wenden uns zu der Frage nach der Sicherheit der numerischen Elemente.

Um die aufgeworfene Frage verständlich zu machen, betrachten wir einen beobachteten K.-G., dessen Glieder in der Urliste nach Zufall variieren. Da nun einerseits bei einem wirklich beobachteten K.-G. der Umfang notwendig endlich bleibt, und da andererseits eine vollständige Ausgleichung des Zufalls nur bei einem unendlich großen Umfange zu erwarten ist, so sind in der Verteilungskurve und ebenso in den daraus hergeleiteten Werten der numerischen Elemente noch Reste von unausgeglichenen Zufälligkeiten enthalten, d. h. man erhält für die genannten Elemente nicht die *Sollwerte,* die bei einem unendlichen Umfange des K.-G. zum Vorschein kommen würden. Kennt man irgendwoher die Sollwerte, so lehrt die Differenz *Soll minus Ist* sofort auch den Einfluß kennen, den der Zufall noch ausgeübt hat; das trifft z. B. bei gewissen Glücksspielen und ähnlichen Versuchen zu, bei denen man die Versuchsbedingungen vollständig in der Hand hat. Weit häufiger bleiben jedoch die Sollwerte der Elemente unbekannt, und das Gleiche gilt dann natürlich auch von den Differenzen „Soll minus Ist“. In diesem Falle muß man offenbar die tatsächlich vorhandenen, aber unbekannt bleibenden Abweichungen von dem Soll als einen unvermeidlichen Mangel des vorgelegten Beobachtungsmaterials hinnehmen. Nun werden die schädlichen Folgen eines solchen Mangels je nach den Umständen von sehr verschiedenem Gewichte sein, und man kann danach die Frage aufwerfen, ob sich nicht

für die zu befürchtenden Nachteile eine brauchbare Maßbestimmung aufstellen lasse, nach der man ziffermäßig das Vertrauen abzuschätzen vermag, das den gefundenen Zahlenwerten der Elemente zuzusprechen ist. Eine derartige Maßbestimmung wird namentlich dann von Wichtigkeit sein, wenn die Untersuchung eines K.-G. nicht nur die Feststellung seines tatsächlichen Verlaufes, sondern auch die Grundlage für weiter gehende Schlüsse zu liefern hat.

Bei der gesuchten Maßbestimmung kann es sich selbstverständlich nicht darum handeln, die tatsächlichen Beträge der Abweichungen von den Sollwerten abzuschätzen, da ja letztere in dem vorliegenden Falle notwendig unbekannt bleiben; man wird sich vielmehr darauf beschränken müssen, anzugeben, welche Abweichungen „im allgemeinen" oder „im Durchschnitt" — oder wie man sonst sagen will — zu *befürchten* sind. Hierzu sollen uns gewisse Streuungsgrößen dienen, deren Entstehung zunächst an einem Beispiele erläutert werden möge.

§ 123. Man denke sich zehn gewöhnliche Würfel gleichzeitig geworfen und die rH., mit der die Eins vorkommt, als beobachtetes x notiert, wobei x offenbar einen der Werte

$$0.0, \quad 0.1, \quad 0.2, \; \ldots \; 1.0$$

annehmen kann. Weiter denken wir uns den Wurf wiederholt und im ganzen m-mal ausgeführt, so daß wir einen aus m Würfen bestehenden „Versuch" erhalten. Diesen Versuch betrachten wir als einen K.-G., dessen Glieder die einzelnen Würfe und dessen Argument die notierten x bilden. Dann gehört dazu ein System bestimmter Werte der Elemente, die wir uns nach irgend einem zulässigen Verfahren aus den beobachteten x hergeleitet denken und zur Unterscheidung von den Sollwerten ebenfalls kurz als *beobachtete* Werte bezeichnen wollen.

Greifen wir aus dem System der Elemente ein bestimmtes Element E heraus und denken uns den angegebenen Versuch unbeschränkt oft wiederholt, so liefert jeder neue Versuch einen neuen beobachteten Wert des betrachteten E, der, wie es der Zufall fügt, teils mit den übrigen Werten von E übereinstimmt, teils von ihnen abweicht. Diese Reihe der E-Werte wollen wir nun als einen neuen K.-G. ansehen, den wir kurz mit $K(E)$ bezeichnen, wobei das Argument von den E-Werten selber gebildet wird. Dann gehört dazu ein gewisses System von numerischen Elementen, von denen hier im besonderen die Streuung in Betracht kommen wird.

Die vorstehende, für den Fall des Würfelspiels angestellte Betrachtung übertragen wir jetzt auf andere K.-G., bei denen erstlich die Glieder in der Urliste nach Zufall variieren, und zweitens die beliebig häufige Wiederholung der Beobachtung des K.-G. wenigstens denkbar ist. Man hat es dann jedesmal zu tun mit einer beobachteten

Kollektivreihe K und den dazu gehörigen beobachteten Elementwerten, ferner ist zu jedem Element E eine gewisse Kollektivreihe $K(E)$ nebst der entsprechenden Streuung str (E) hinzuzudenken. Nun lehren die Zahlen des in § 91 berechneten Täfelchens, daß die Streuung eines K.-G. eine summarische Vorstellung von der Ausbreitung seines Arguments gibt; je kleiner die Streuung ausfällt, desto enger drängen sich die Argumente um ihren Mittelwert zusammen. Danach werden wir von einem Elementwert E, der aus dem vorgelegten K abgeleitet worden ist, folgendes aussagen dürfen: je kleiner str (E), desto eher können wir erwarten, daß für das beobachtete E die Differenz „Soll minus Ist" unterhalb einer vorgeschriebenen Grenze liegen werde. Mit Rücksicht hierauf wählen wir str (E) als *Maß für die Unsicherheit*, die dem beobachteten E zuzusprechen ist.

Die eingeführte Maßgröße str (E) mißt nicht den Betrag der Differenz „Soll minus Ist", der bei dem beobachteten E tatsächlich stattfindet, sondern nur den durchschnittlichen absoluten Betrag dieser Differenz. Daher kann das Vertrauen, das man auf Grund des vorliegenden Wertes von str (E) in das gefundene E zu setzen geneigt ist, unter Umständen auch getäuscht werden: wenn z. B. für ein Element E zwei Bestimmungen mit verschiedenen Werten von str (E) vorliegen, so ist es sehr wohl denkbar, daß die Bestimmung mit der kleineren Streuung die stärkere Abweichung von dem Sollwerte besitzt. Dadurch wird indessen nichts an dem Satze geändert, daß, wie das Täfelchen in § 91 lehrt, *der Regel nach* dem kleineren str (E) auch die größere Vertrauenswürdigkeit entspricht. Um ein Gleichnis zu gebrauchen denke man sich einen geschickten Schützen, dem zwei verschiedene Gewehre, nämlich eine moderne Präzisionswaffe und eine altmodische Luntenflinte zur Verfügung stehen. Handelt es sich um je einen Schuß aus jedem Gewehr, so kann es der Zufall sehr wohl so fügen, daß der Schütze mit der mangelhaften Waffe trifft, mit der anderen dagegen fehlt. Trotzdem ist niemand darüber in Zweifel, daß das moderne Gewehr dem alten für den regelmäßigen Gebrauch vorzuziehen ist, und daß ihre Verschiedenheit in den Trefferbildern einer größeren Reihe von Schüssen deutlich zum Ausdruck kommen würde.

Gegen die gewählte Maßbestimmung läßt sich der Einwand erheben, daß sie nicht frei von Willkür sei, und daß man statt str (E) füglich mit demselben Rechte irgend einen anderen Mittelwert aus den absoluten Abweichungen der E von ihrem Durchschnitt zugrunde legen könnte. Der Einwand ist vollkommen richtig, trifft jedoch *jeden* Versuch einer derartigen Sicherheitsschätzung. Denn als Grundlage einer solchen Schätzung müßte immer die durch $K(E)$ gegebene Verteilung der E dienen, und diese hängt außer von str (E) auch noch von anderen Parametern, namentlich, wie das erwähnte Täfelchen lehrt,

von den geraden D-Koeffizienten der Reihe $K(E)$ ab; es ist also nicht möglich, die Ausbreitung der E allgemein in *erschöpfender* Weise durch eine einzige Zahl zu charakterisieren, wie das doch zur Aufstellung eines unbedingt einwandfreien Unsicherheitsmaßes nötig sein würde. Aus diesem Grunde wurde schon oben nur von einer *summarischen* Darstellung der Ausbreitung der E gesprochen. Sobald aber nur eine summarische Schätzung in Frage kommt, steht uns die Wahl zwischen verschiedenen Unsicherheitsmaßen frei, und wir dürfen dann äußere Zweckmäßigkeitsgründe den Ausschlag für $\mathrm{str}(E)$ geben lassen.

§ 124. Da bei einer vorgelegten Kollektivreihe K die Verteilung $K(E)$ eines zu K gehörigen numerischen Elements E davon abhängt, wie sich E aus den beobachteten Argumentwerten von K zusammensetzt, so haben wir jetzt anzugeben, wie die Werte der E zu berechnen sind. Diese Frage läßt sich für ein unstetiges K ohne weiteres beantworten. Denn einerseits stellen sich die E als die Durchschnitte gewisser ganzer Funktionen des Arguments x, nämlich der Verbindungen

$$x, \quad (x-c)^2, \quad \mathfrak{R}(hx-hc)_n$$

dar, andererseits besteht für ein unstetiges K mit der Verteilung $\mathfrak{U}(x)$ die Gleichung

$$\mathfrak{D}[T(x)] = \sum T(x)\mathfrak{U}(x), \tag{1}$$

wo die Summe über die möglichen x zu erstrecken ist; man erhält also das gesuchte beobachtete E, wenn man für das zugehörige $T(x)$ in (1) die Produktsumme auf der rechten Seite mit den beobachteten x und den entsprechenden beobachteten Werten von $\mathfrak{U}(x)$ bildet. Diese Bemerkung wollen wir noch etwas anders einkleiden. Wir denken uns die beobachteten x in der Ordnung, in der sie durch die Urliste gegeben werden, fortlaufend mit den Nummern $1, 2, \ldots m$ versehen, wo m den Umfang des vorgelegten K.-G. bedeutet, ferner werde der in der angegebenen Weise berechnete Durchschnitt zum Unterschiede von dem Sollwerte durch das Symbol Δ bezeichnet, dann hat man, wenn für den reziproken Wert von m der Buchstabe M geschrieben wird,

$$\Delta[T(x)] = \sum_r MT(x_r), \quad (r = 1, 2, \cdots m). \tag{2}$$

Um die Richtigkeit dieser Gleichung einzusehen, hat man sich nur zu vergegenwärtigen, wie die in (1) zugrunde gelegte primäre Verteilungstafel aus der in (2) benutzten Urliste entsteht.

Die Formel (2) besitzt den Vorteil, daß sie auch noch bei den stetigen K.-G. ohne weiteres angewendet werden darf, wenn wir die Werte der x als abrundungsfrei beobachtet voraussetzen. In Wirklichkeit trifft diese Voraussetzung allerdings nicht zu, daher wird sich

der Einfluß der Abrundung in dem Werte von $\Delta(T)$ geltend machen, und zwar im allgemeinen um so stärker, je größer die vorgenommene Abrundung ist, oder je länger die zwischen den Wechselpunkten liegenden Teilstrecken genommen werden. Da indessen die Abrundung mit dem Wesen eines stetigen K.-G. an sich nichts zu tun hat, so wollen wir vorläufig keine Rücksicht darauf nehmen und die stetigen K.-G. ebenfalls nach (2) behandeln, mit dem Vorbehalt, daß wir den vernachlässigten Einfluß der Abrundung nachträglich noch besonders untersuchen.

Die in (2) enthaltene Vorschrift zur Berechnung der beobachteten Werte von $\mathfrak{D}(T)$ können wir als *direkte Mittelbildung* bezeichnen; sie besitzt offenbar den Vorzug, daß sie keinerlei willkürliche Ansätze heranzieht und die Beobachtungen unmittelbar in der Gestalt benutzt, in der sie gegeben zu denken sind. Liegt ein unstetiger K.-G. mit einer nur mäßigen Anzahl beobachteter $\mathfrak{U}(x)$ vor, so liefert das direkte Mittel einen auch praktisch durchaus brauchbaren Rechnungsgang zur Herstellung der gesuchten Φ-Reihe. Ist dagegen die Anzahl der $\mathfrak{U}(x)$ sehr groß oder hat man es mit einem stetigen K.-G. zu tun, so würde die direkte Rechnung in der hier vorausgesetzten Gestalt allerdings beschwerlich ausfallen, und man wird sich daher bei solchen Reihen noch nach anderen Wegen umsehen. Wir brauchen indessen auf diesen Punkt im Augenblicke nicht näher einzugehen, weil man an den Rechnungsgang, wie er auch eingerichtet sein möge, die Forderung zu stellen hat, daß sein Ergebnis nicht wesentlich von den Zahlen der direkten Mittelbildung abweiche.

§ 125. Bei der vorgelegten Kollektivreihe K mit der Verteilung $\mathfrak{V}(x)$ oder $\mathfrak{U}(x)$ erscheint jedes berechnete Element E als eine bestimmte, durch die Gleichung (2) definierte Funktion der Größen x_1, $x_2, \ldots x_m$, die der Reihenfolge ihrer Nummern entsprechend das erste, zweite usw. Argument der Urliste bedeuten. Nun wird hier von K vorausgesetzt, daß bei unbeschränkt wiederholter Beobachtung von K jedes zu einer vorgeschriebenen Nummer gehörige x — sagen wir z. B. x_1 — mit einer Verteilung auftrete, welche durch die zu K gehörige Verteilungsfunktion $\mathfrak{V}(x)$ oder $\mathfrak{U}(x)$ geregelt ist, und daß ferner dieses Auftreten von x_1 ganz unabhängig von den gleichzeitigen Werten der übrigen x_k vor sich gehe. Damit gelangen wir aber zu den Voraussetzungen, die bei der Untersuchung der Argumentmischung zugrunde gelegt wurden, so daß die dort entwickelten Sätze über die Verteilung von Systemen simultaner Argumentwerte unmittelbar Anwendung finden können. Der ganze Unterschied gegen früher besteht nur darin, daß die früher mit $U(x_k)$ bezeichneten und im allgemeinen voneinander verschiedenen Verteilungsgrößen jetzt sämtlich die gleiche Gestalt besitzen und durch den Ausdruck $\mathfrak{V}(x)dx$ oder $\mathfrak{U}(x)$ bestimmt sind. Des weiteren gilt auch die frühere Bemerkung über den Ge-

brauch des $\mathfrak{D}$-Zeichens, das sich immer auf die Verteilung der hinter dem Zeichen stehenden Argumente bezieht.

Mit Rücksicht auf die vorstehenden Bemerkungen folgen für die hier betrachtete Kollektivreihe K aus der Relation

$$E = T(x_1, x_2, \cdots x_m)$$

die Gleichungen

$$\mathfrak{D}(E) = \mathfrak{D}(T), \qquad \text{str}(E) = \text{str}(T),$$

$$\text{str}(T)^2 = \mathfrak{D}[[T - \mathfrak{D}(T)]^2]$$

oder

$$\text{str}(T)^2 = \mathfrak{D}(T^2) - \mathfrak{D}[2\,T\mathfrak{D}(T)] + \mathfrak{D}(T)^2.$$

Nun ist in dem mittleren Gliede rechts das in [ ] stehende $\mathfrak{D}(T)$ für die vor [ ] stehende $\mathfrak{D}$-Operation eine Konstante, kann also herausgezogen werden, so daß man nach der erforderlichen Zusammenziehung

$$\text{str}(E)^2 = \text{str}(T)^2 = \mathfrak{D}(T^2) - \mathfrak{D}(T)^2 \tag{3}$$

erhält. Besitzt E die Gestalt $aT + b$, wo a und b Konstanten bedeuten, so folgt aus den in § 96—97 entwickelten Sätzen über die linearen Transformationen

$$\text{str}(E) = \pm\, a\; \text{str}(T), \tag{4}$$

wo $\pm\, a$ stets positiv sein muß. Dies festgestellt setzen wir für das vorgelegte K die Sollwerte der Elemente nach den Gleichungen

$$c = \mathfrak{D}(x), \qquad s = \mathfrak{D}[(x-c)^2], \qquad h^2 = 1 : 2s^2, \tag{5}$$

$$D(c, h)_n = \mathfrak{D}[\mathfrak{R}(hx - hc)_n] \tag{6}$$

an, denen wir noch die mit einem beliebigen nicht normalen Parameterpaar e, k berechneten Koeffizienten

$$D(e, k)_n = \mathfrak{D}[\mathfrak{R}(kx - ke)_n] \tag{7}$$

zuordnen. Entsprechend setzen wir für die „beobachteten“ Elementwerte die vermittelst der Operation

$$\Delta[f(x)] = \sum_r Mf(x_r), \qquad (r = 1, 2, \cdots m) \tag{8}$$

gebildeten Gleichungen

$$C = \Delta(x) = \sum_r Mx_r, \tag{9}$$

$$S^2 = \Delta[(x-C)^2] = \sum_r M(x_r - C)^2, \qquad H^2 = 1 : 2S^2, \tag{10}$$

$$\Delta(C, H)_n = \Delta[\mathfrak{R}(Hx - HC)_n] = \sum_r M\mathfrak{R}(Hx_r - HC)_n, \tag{11}$$

$$\Delta(e, k)_n \;= \Delta[\mathfrak{R}(kx - ke)_n] \;= \sum_r M\mathfrak{R}(kx_r - ke)_n \tag{12}$$

an.

§ 126. Die Ausdrücke (9) bis (12) hat man nun der Reihe nach für die vorhin mit T bezeichnete Funktion der x_r zu setzen und den Wert von $\mathrm{str}(T)$ aufzusuchen. Zu dem Ende sind jedesmal die Durchschnitte $\mathfrak{D}(T)$ und $\mathfrak{D}(T^2)$ zu bilden und nach (3) miteinander zu verbinden. Hierbei darf man Durchschnitte von der Form $\mathfrak{D}[f(x_r)]$, die nach dem Argument x_r zu bilden sind, einfach durch $\mathfrak{D}[f(x)]$ ersetzen, da ja das Argument x_r dieselbe Verteilung besitzt, wie das Argument der vorgelegten Kollektivreihe K.

Dies vorausgeschickt führen wir noch die Abkürzungen

$$x - c = y, \qquad x_r - c = y_r, \qquad C - c = F$$

ein, womit der Reihe nach

$$\mathfrak{D}(y_r) = \mathfrak{D}(y) = \mathfrak{D}(x - c) = \mathfrak{D}(x) - c = 0, \tag{13}$$

$$\mathfrak{D}(y_r^2) = \mathfrak{D}(y^2) = s^2, \tag{14}$$

$$\Delta(y) = \sum_r M y_r = \sum_r M(x_r - c) = C - c = F \tag{15}$$

erhalten wird. Bildet man jetzt den Durchschnitt von F, so wird nach (15) und (13)

$$\mathfrak{D}(F) = \mathfrak{D}[\Delta(y)] = \mathfrak{D}[\sum_r M y_r] = \sum_r M \mathfrak{D}(y_r) = 0,$$

woraus wegen $F = C - c$

$$0 = \mathfrak{D}(F) = \mathfrak{D}(C - c) = \mathfrak{D}(C) - c,$$

$$\mathfrak{D}(C) = c \tag{16}$$

folgt. Ferner ist

$$\mathfrak{D}(F^2) = \mathfrak{D}[(\sum_r M y_r)^2] = \mathfrak{D}[\sum_q \sum_r M^2 y_q y_r].$$

Für ungleiche Indizes wird nun unter Berücksichtigung von (13)

$$\mathfrak{D}(y_q y_r) = \mathfrak{D}(y_q)\mathfrak{D}(y_r) = 0,$$

so daß nur die Fälle $q = r$ zu berücksichtigen sind, woraus wegen (14)

$$\mathfrak{D}(F^2) = \sum_r M^2 \mathfrak{D}(y_r^2) = M s^2 \tag{17}$$

erhalten wird. Da $\mathfrak{D}(F)$ verschwindet, so ist wegen (3)

$$\mathrm{str}(F)^2 = M s^2, \tag{18}$$

und weiter, da sich F und C nur um eine Konstante unterscheiden,

$$\mathrm{str}(C)^2 = M s^2. \tag{19}$$

Hiermit ist die Streuung für den beobachteten Wert von $\mathfrak{D}(x)$ gefunden.

§ 127. Bei der weiteren Rechnung wollen wir das Summenzeichen nur einfach und ohne angefügte Indizes schreiben, wenn die Summa-

tion über alle hinter dem Zeichen stehenden Indizes auszudehnen ist. Zur Berechnung von S hat man dann zunächst nach (10)

$$S^2 = \sum M(x_r - C)^2 = \sum M(y_r - F)^2$$
$$= \sum My_r^2 - 2F\sum My_r + F^2,$$

also wegen (15)

$$S^2 = \sum My_r^2 - F^2. \tag{20}$$

Daraus folgt unter Berücksichtigung von (14) und (17)

$$\mathfrak{D}(S^2) = \sum M\mathfrak{D}(y_r^2) - \mathfrak{D}(F^2) = s^2 - Ms^2,$$

oder mit Einführung der Abkürzung $N = 1 - M$

$$\mathfrak{D}(S^2) = Ns^2. \tag{21}$$

Zur Berechnung der Streuung von S^2 haben wir noch den Durchschnitt der Größe S^4 zu bilden, die nach (20) zunächst in der Gestalt

$$S^4 = \sum M^2 y_q^2 y_r^2 - 2F^2 \sum My_r^2 + F^4$$

erscheint. Bezeichnet man für den Augenblick die Durchschnitte der drei Bestandteile auf der rechten Seite mit I, II, III, so wird zunächst

$$\mathrm{I} = \sum M^2 \mathfrak{D}(y_q^2 y_r^2).$$

Trennt man die Fälle „q gleich r" und „q ungleich r", so erhält man für I die Beiträge

$$\sum M^2 \mathfrak{D}(y_r^4) \quad \text{und} \quad \sum M^2 \mathfrak{D}(y_q^2) \cdot \mathfrak{D}(y_r^2),$$

woraus

$$\mathrm{I} = M\mathfrak{D}(y^4) + m(m-1)M^2 s^4$$

folgt. Ferner erhält man für II wegen (15) zunächst

$$\mathrm{II} = -2\mathfrak{D}[F^2 \sum My_r^2] = -2\sum M^3 \mathfrak{D}(y_r^2 y_p y_q).$$

Sind p und q ungleich, so kann r höchstens mit einer der beiden Nummern p, q übereinstimmen, und man darf in diesem Falle für die zuletzt stehende D-Größe

$$\text{entweder} \quad \mathfrak{D}(y_r^2 y_p) \cdot \mathfrak{D}(y_q) \quad \text{oder} \quad \mathfrak{D}(y_r^2 y_q) \cdot \mathfrak{D}(y_p)$$

schreiben, erhält also jedesmal den Wert Null. Demnach wird, weil nur noch die Fälle $p = q$ zu berücksichtigen sind,

$$\mathrm{II} = -2\sum M^3 D(y_q^2 y_r^2) = -2M \cdot \mathrm{I}.$$

Schließlich ist noch

$$\mathrm{III} = \mathfrak{D}(F^4) = \sum M^4 \mathfrak{D}(y_n y_p y_q y_r)$$

zu bilden. Hierin sind die Summenglieder, wie man beim Durchgehen der möglichen Fälle erkennt, nur dann von Null verschieden, wenn einer der vier Fälle

$$1)\ n = p = q = r, \qquad 2)\ n = p, \quad q = r,$$
$$3)\ n = q, \quad p = r, \qquad 4)\ n = r, \quad p = q$$

vorliegt, wobei in 2), 3), 4) der Fall 1) auszuschließen ist. Daraus ergibt sich mit Unterdrückung der Zwischenrechnung

$$\mathrm{III} = M^3\mathfrak{D}(y^4) + 3m(m-1)M^4s^4.$$

Faßt man nunmehr I, II, III zusammen, so wird nach einer kleinen Reduktion

$$\begin{aligned}\mathfrak{D}(S^4) &= (1-2M)[M\mathfrak{D}(y^4) + (m-1)Ms^4] + M^3\mathfrak{D}(y^4) + 3(m-1)M^3s^4\\ &= MN^2\mathfrak{D}(y^4) + (N^3 + 2M^2N)s^4,\end{aligned}$$

woraus unter Beachtung von (3) und (21)

$$\begin{aligned}\mathrm{str}\,(S^2)^2 &= \mathfrak{D}(S^4) - \mathfrak{D}(S^2)^2\\ &= MN^2\mathfrak{D}(y^4) + (2M^2N - MN^2)s^4 \qquad (22)\end{aligned}$$

folgt. Dieser Ausdruck läßt sich noch etwas umformen. Da nämlich

$$96\,\mathfrak{R}(hy)_4 = 4(hy)^4 - 12(hy)^2 + 3$$

ist, so wird, wenn man den Durchschnitt bildet und die Beziehung $1 = 2h^2s^2$ beachtet,

$$\begin{aligned}96\,D(c,h)_4 &= 4h^4\mathfrak{D}(y^4) - 12h^2s^2 + 3 = 4h^4\mathfrak{D}(y^4) - 3,\\ 96\,s^4D(c,h)_4 &= \mathfrak{D}(y^4) - 3\,s^4.\end{aligned}$$

Benutzt man diese Gleichung, um in (22) die Größe $\mathfrak{D}(y^4)$ herauszuschaffen, so erhält man

$$\mathrm{str}\,(S^2)^2 = 2MNs^4(1 + 48ND(c,h)_4). \qquad (23)$$

Da in der vorstehenden Gleichung die rechte Seite stets positiv sein muß, so folgt, daß bei den hier betrachteten K.-G. der Koeffizient D_4 stets oberhalb des reziproken Wertes von $-48N$ liegen muß.

Die Formeln (19) und (23) setzen, wenn man darnach die gesuchten Streuungen berechnen will, voraus, daß man die Sollwerte der Elemente kennt, während in Wirklichkeit nur die beobachteten Werte bekannt sind. Dieser Mangel läßt sich natürlich nicht beseitigen, wohl aber durch einen in der Ausgleichungsrechnung seit langer Zeit benutzten Kunstgriff mildern. Man ersetzt nämlich nach (21) den Sollwert von s^2 zunächst durch den Quotienten $\mathfrak{D}(S^2) : N$ und nimmt hierin für den unbekannten Durchschnitt $\mathfrak{D}(S^2)$ den beobachteten Wert von S^2. Dadurch berücksichtigt man einigermaßen den Umstand, daß S^2 nach Ausweis von (21) im *Durchschnitt* kleiner als sein Sollwert gefunden wird, da ja N kleiner als Eins ist.

Die Formel (23) liefert die Streuung von S^2, während man eigentlich die Streuung von S haben möchte. Da indessen S eine irrationale Funktion der x_r ist, so muß man im allgemeinen darauf verzichten, für $\mathrm{str}\,(S)$ einen genügend einfachen und zur numerischen Rechnung brauchbaren Ausdruck aufzustellen. Überdies läßt sich eine summa-

rische Schätzung der Unsicherheit von S auch auf Grund des Wertes von $\operatorname{str}(S^2)$ erreichen.

§ 128. Versucht man die Koeffizienten $\Delta(C, H)_n$ in ähnlicher Weise wie die Größen C und S zu behandeln, so stellt sich heraus, daß dies nicht angeht, weil jene Koeffizienten nicht mehr ganze rationale Funktionen der x_r sind. Setzt man statt dessen die Größen

$$S^n \Delta(C, H)_n$$

an, so kommt man allerdings wieder auf ganze rationale Ausdrücke, indessen werden selbst für die niedrigsten Nummern $n = 3$ und $n = 4$ die Formeln nicht sonderlich einfach. Aus diesem Grunde wollen wir uns auf die mit den Parametern e, k gebildeten Größen beschränken, wobei wir zur Abkürzung

$$D_n = D(e, k)_n, \qquad \Delta_n = \Delta(e, k)_n$$

schreiben. Setzt man noch

$$u_r = k(x_r - e), \qquad u = k(x - e),$$

so wird

$$\Delta_n = \sum_r M \mathfrak{R}(u_r)_n, \qquad \mathfrak{D}(\Delta_n) = \sum_r M \mathfrak{D}[\mathfrak{R}(u_r)_n] = D_n, \tag{24}$$

$$\Delta_n^2 = \sum_q \sum_r M^2 \mathfrak{R}(u_q)_n \mathfrak{R}(u_r)_n.$$

Nimmt man von der letzten Gleichung den Durchschnitt und trennt wieder die Fälle mit gleichen und ungleichen q, r, so wird nach der nötigen Reduktion

$$\mathfrak{D}(\Delta_n^2) = M \mathfrak{D}\{[\mathfrak{R}(u)_n]^2\} + N D_n^2,$$

woraus mit Rücksicht auf (24)

$$\operatorname{str}(\Delta_n)^2 = M \mathfrak{D}\{[\mathfrak{R}(u)_n]^2\} - M D_n^2$$

folgt. Um die hierin auftretenden Quadrate der $\mathfrak{R}$-Größen durch die $\mathfrak{R}$ selber linear auszudrücken, setzen wir zunächst die beiden Gleichungen

$$\sum \mathfrak{R}(u)_p (2v)^p = \exp(-2uv - v^2), \qquad \sum \mathfrak{R}(u)_q (2w)^q = \exp(-2uw - w^2)$$

an und erhalten aus ihrem Produkt

$$\begin{aligned}\sum \mathfrak{R}(u)_p \mathfrak{R}(u)_q (2v)^p (2w)^q &= \exp[-2u(v+w) - (v+w)^2 + 2vw] \\ &= \exp(2vw) \sum_r \mathfrak{R}(u)_r (2v + 2w)^r.\end{aligned}$$

Setzt man rechter Hand die Reihen

$$\exp(2vw) = \sum \frac{1}{s!}(2vw)^s, \qquad (2v + 2w)^r = \sum_t \frac{r!}{t!(r-t)!}(2v)^t (2w)^{r-t}$$

ein, so wird zunächst

$$\sum \mathfrak{R}(u)_p \mathfrak{R}(u)_q (2v)^p (2w)^q = \sum \frac{r!}{s!\,t!\,(r-t)!} 2^{-s} \mathfrak{R}(u)_r (2v)^{s+t} (2w)^{r+s-t}.$$

Hieraus ergibt sich der Ausdruck für das Produkt aus $\mathfrak{R}_p$ und $\mathfrak{R}_q$, wenn man in der Summe

$$\sum \frac{r!}{s!\,t!\,(r-t)!}\, 2^{-s}\mathfrak{R}(u)_r$$

sich auf diejenigen Werte von r, s, t beschränkt, die den Bedingungen

$$s+t=p, \qquad r+s-t=q$$

genügen. Ersetzt man demgemäß in der vorstehenden Summe die Indizes r und t durch

$$p+q-2s \quad \text{und} \quad p-s,$$

so wird

$$\mathfrak{R}(u)_p\,\mathfrak{R}(u)_q = \sum_s \frac{(p+q-2s)!}{s!\,(p-s)!\,(q-s)!}\, 2^{-s}\mathfrak{R}(u)_{p+q-2s}, \tag{25}$$

wobei s von Null an bis zu der kleineren der beiden Zahlen p und q zu laufen hat. Für $p=q$ erhält man dann

$$[\mathfrak{R}(u)_p]^2 = \sum_s \frac{(2p-2s)!}{s!\,(p-s)!\,(p-s)!}\, 2^{-s}\mathfrak{R}(u)_{2p-2s}$$

oder mit $r=p-s$

$$[\mathfrak{R}(u)_p]^2 = \sum_r \frac{(2r)!}{(p-r)!\,r!\,r!}\, 2^{r-p}\mathfrak{R}(u)_{2r},$$

worin r von 0 bis p zu laufen hat.

Nimmt man von der letzten Gleichung den Durchschnitt, so ergibt sich für die gesuchte Streuung

$$\operatorname{str}(\Delta_n)^2 = M\sum_r \frac{(2r)!}{(n-r)!\,r!\,r!}\, 2^{r-n}D_{2r} - MD_n^2. \tag{26.a}$$

Auch diese Formel wird man nur in beschränktem Maße anwenden, weil sie die Kenntnis der höheren D_n voraussetzt, und weil diese Koeffizienten zum Teil mit großen Zahlenfaktoren verbunden auftreten. So ist z. B. für $n=3$

$$48\operatorname{str}(\Delta_3)^2 = M(1+12D_2+144D_4+96D_6-D_3^2). \tag{26.b}$$

Ein einfaches Ergebnis erhält man eigentlich nur, wenn das vorgelegte K merklich dem E.-G. folgt, denn dann verschwinden, von D_0 abgesehen, die auftretenden D-Größen für $e=c$ und $k=h$, und man erhält

$$\operatorname{str}[\Delta(c,h)_n]^2 = M : 2^n n!. \tag{26.c}$$

Man kann diese Formel dazu benutzen, um zu beurteilen, ob ein vorgelegtes K in merklicher Weise von dem E.-G. abweicht, denn bei einem K, das dem E.-G. folgt, müssen die beobachteten $\Delta(c,h)_n$ von derselben Größenordnung wie ihre Streuungen m, aber nicht wesentlich größer, sein.

§ 129. Faßt man die vorstehenden Entwicklungen zusammen, so sieht man zunächst, daß die gefundenen Streuungsgrößen sämtlich die

Quadratwurzel von M oder $1:m$ als Faktor enthalten, also unter sonst gleichen Umständen umgekehrt proportional zu der Wurzel aus dem Umfange m des vorgelegten K.-G. variieren. Ferner lassen sich die Streuungen von C und S^2 nach (19) und (23) ohne Schwierigkeit berechnen. Hingegen sind die Formeln für die Streuungen der beobachteten D-Koeffizienten zum allgemeinen Gebrauche weniger geeignet, namentlich wenn man beachtet, daß in den Δ_n, die man ja statt der unbekannten D_n einsetzen muß, die Zufälligkeiten der Beobachtung in der Regel einen um so stärkeren Einfluß ausüben, je höher die Ordnungsnummer n ist. Allerdings gibt es noch einen anderen Weg, auf dem man die Streuung eines Δ_n ohne die Kenntnis der höheren Δ-Größen ermitteln kann, jedoch würden wir damit das Gebiet der sogenannten Ausgleichungsrechnung betreten und die für die vorliegende Darstellung gesteckten Grenzen überschreiten müssen.

Da die Ausdrücke (26.a) ihrer Bedeutung wegen niemals negativ ausfallen dürfen, so liefert jeder Ausdruck dieser Art eine Ungleichung, der die D-Koeffizienten der hier betrachteten K.-G. zu genügen haben. Es wäre eine ganz interessante, hier jedoch zu weit führende Aufgabe, das vollständige Formensystem solcher Ungleichungen aufzusuchen, die offenbar als Kriterien dafür dienen können, ob in der Urliste eines vorgelegten K.-G. zwischen den einzelnen Gliedern Sukzessionsabhängigkeit besteht oder nicht.

§ 130. Außer den numerischen Elementen werden auch die beobachteten Werte von $\mathfrak{S}(x)$, $\mathfrak{B}(x)$ und $\mathfrak{U}(x)$ durch die der Beobachtung anhaftenden Zufälligkeiten beeinflußt, so daß wir auch für diese Größen noch die Streuungen aufzusuchen haben.

Führt man wieder das Zeichen

$$2E(X, x) = \operatorname{sg}(X - x) + 1$$

ein, so ist der Sollwert von $\mathfrak{S}(x)$ durch

$$\mathfrak{S}(X) = \mathfrak{D}[E(X, x)]$$

gegeben, da ja die Größe E gleich Eins oder Null ist, je nachdem x unterhalb oder oberhalb X liegt. Ebenso ergibt sich der beobachtete Wert von $\mathfrak{S}(X)$, den wir mit $\mathfrak{T}(X)$ bezeichnen wollen, aus

$$\mathfrak{T}(X) = \Delta[E(X, x)] = \sum ME(X, x_r).$$

Nimmt man hiervon den Durchschnitt nach allen Wertsystemen der x_r, so wird

$$\mathfrak{D}[\mathfrak{T}(X)] = \mathfrak{S}(X). \tag{27}$$

Bildet man weiter von

$$\mathfrak{T}(X)^2 = \sum M^2 E(X, x_\varrho) \cdot E(X, x_r)$$

den Durchschnitt und trennt die Fälle mit gleichen und ungleichen q, r, so erhält man für ungleiche q, r

$$\mathfrak{D}[E(X, x_q)\, E(X, x_r)] = \mathfrak{D}[E(X, x_q)] \cdot \mathfrak{D}[E(X, x_r)] = \mathfrak{S}(X)^2.$$

Ferner ist, da für die E-Größen nur die Werte Null oder Eins in Betracht kommen,

$$E(X, x_r)^2 = E(X, x_r),$$

woraus

$$\mathfrak{D}[E(X, x_r)^2] = \mathfrak{S}(X)$$

folgt. Damit ergibt sich

$$\mathfrak{D}(\mathfrak{T}^2) = M\mathfrak{S}(X) + m(m-1)M^2\mathfrak{S}(X)^2,$$

und weiter nach der Formel (3)

$$\text{str}\,[\mathfrak{T}(X)]^2 = M\mathfrak{S}(X)[1 - \mathfrak{S}(X)]. \tag{28}$$

Aus der vorstehenden Gleichung folgt noch für die Streuung des relativen Überschusses der Ausdruck

$$\text{str}\,[2\,\mathfrak{T}(X) - 1]^2 = 4M\mathfrak{S}(X)[1 - \mathfrak{S}(X)]. \tag{29}$$

Will man statt der Summenwerte die Verteilungswerte $\mathfrak{B}$ oder $\mathfrak{U}$ untersuchen, so setze man mit der Annahme $X > Y$

$$S(X, Y) = \mathfrak{S}(X) - \mathfrak{S}(Y) = \mathfrak{D}[E(X, x) - E(Y, x)],$$
$$T(X, Y) = \mathfrak{T}(X) - \mathfrak{T}(Y) = \sum M[E(X, x_r) - E(Y, x_r)].$$

Dann ist zunächst

$$\mathfrak{D}[T(X, Y)] = S(X, Y). \tag{30}$$

Ferner wird

$$T(X, Y)^2 = \sum M^2[E(X, x_q) - E(Y, x_q)] \cdot [E(X, x_r) - E(Y, x_r)],$$

woraus man ähnlich wie oben

$$\mathfrak{D}[T(X, Y)^2] = M\mathfrak{D}\{[E(X, x) - E(Y, x)]^2\} + NS(X, Y)^2$$

erhält. Da nun die Differenz $E(X, x) - E(Y, x)$ dieselben Werte wie ihr Quadrat besitzt, so liefert die vorstehende Gleichung mit (30) verbunden

$$\text{str}\,[T(X, Y)]^2 = MS(X, Y)[1 - S(X, Y)]. \tag{31}$$

Die Formeln (28) und (31) sind identisch mit einer Beziehung, die uns später nochmals bei dem *Bernoulli*schen Satze begegnen wird.

Zum Schlusse möge nochmals darauf hingewiesen werden, daß die Anwendung der hier abgeleiteten Streuungsgrößen als Unsicherheitsmaß an die Bedingung gebunden ist, daß die x_r unabhängig voneinander variieren. Ist diese Bedingung nicht erfüllt, so kann man natürlich immer noch die Streuungen berechnen, man darf sie aber nicht ohne weiteres als Unsicherheitsmaß deuten.

Sechszehnte Vorlesung.

Einfluß der Abrundung.

§ 131. Bei der Untersuchung der Streuungsgrößen in der vorangehenden Vorlesung hatten wir den Einfluß der Abrundung, die bei den Argumenten eines beobachteten stetigen K.-G. auftritt, vorläufig beiseite gelassen, mit dem Vorbehalte, diesen Gegenstand noch besonders zu behandeln. Das soll jetzt geschehen. Zu dem Ende greifen wir auf die Grundformel der direkten Mittelbildung zurück, nämlich auf die Gleichung

$$\Delta[f(x)] = \sum_r Mf(x_r), \quad (r = 1, 2, \ldots m). \qquad (1)$$

Hierin ist M der reziproke Wert des Umfanges m, ferner bedeuten die x_r die einzelnen in der Urliste aufgeführten Argumentwerte. Letztere waren bei der Untersuchung der Streuungsgrößen als abrundungsfrei vorausgesetzt worden, während man, wenn die Rechnung auf Grund der vorgelegten beobachteten Urliste wirklich ausgeführt werden soll, das einzelne x_r jedesmal durch den entsprechenden abgerundeten Wert y_r zu ersetzen hat. Nun wird bei einem beobachteten K.-G. der Wert des einzelnen x_r vom Zufall beeinflußt, und das Gleiche gilt von der Abrundung $x_r - y_r$, so daß es zunächst nicht möglich ist, die Wirkung der Abrundung von der des Zufalls zu trennen und für sich allein genau anzugeben. Will man nun trotzdem von der Wirkung der Abrundung eine deutliche ziffermäßige Vorstellung gewinnen, so hat man offenbar von Verteilungstafeln auszugehen, bei denen die Zufälligkeiten der Beobachtung nicht in Betracht kommen. Das erreicht man, wenn man Tafeln benutzt, die nicht beobachtet, sondern auf Grund einer analytischen Darstellung berechnet worden sind. Denn dann sind die wahren Werte der numerischen Elemente im voraus bekannt und können unmittelbar mit denjenigen Werten verglichen werden, welche aus der berechneten Verteilungstafel mit ihren abgerundeten Argumenten vermittelst der Operation Δ entspringen.

Demgemäß denken wir uns jetzt irgend eine stetige Verteilung $\mathfrak{V}(x)$ nebst der zugehörigen Summenfunktion $\mathfrak{S}(x)$ durch die Darstellung

$$2\mathfrak{S}(x) - 1 = \sum D_p \Phi(u)_p \qquad (2)$$

gegeben, in der die D_p bekannt sind, und in der ferner die irgendwie festgesetzten Parameter c und h des Hilfsarguments

$$u = h(x - c) \qquad (3)$$

nicht der Normalform anzugehören brauchen. Weiter denken wir uns ein System von Wechselpunkten festgesetzt, die wir sogleich als

äquidistant annehmen wollen, weil diese Anordnung bei den stetigen K.-G. die Regel bildet. Damit sind dann sofort auch die Teilstrecken zwischen den benachbarten Wechselpunkten sowie die Halbierungspunkte der Teilstrecken gegeben. Berechnet man nun für die einzelnen Wechselpunkte nach (2) die $\mathfrak{S}(x)$, so erhält man die Summentafel. Subtrahiert man ferner in der Summentafel jedes $\mathfrak{S}(x)$ von dem folgenden, so erhält man die Verteilungszahlen, denen als Argument die Halbierungspunkte der entsprechenden Teilstrecken gegenüber zu stellen sind, um die Verteilungstafel zu bilden. In einer dergestalt konstruierten Tafel sind offenbar alle Unregelmässigkeiten zufälliger Art ausgeschlossen, im übrigen verhält sie sich nicht anders als eine beobachtete Tafel und läßt sich deswegen auch ebenso behandeln.

§ 132. Der Deutlichkeit halber werde das stetig veränderliche Argument wie bisher mit x bezeichnet, das abgerundete Argument der betrachteten Verteilungstafel dagegen mit x' und die zu jedem x' gehörige Verteilungszahl mit $H(x')$, dann sind die x' die Halbierungspunkte der Teilstrecken von x. Denkt man sich ferner — um die Vorstellung zu fixieren — die $H(x')$ mit einer bestimmten Stellenzahl, z. B. mit vier Dezimalen, angesetzt, so läßt sich die Sache so auffassen, als sei die Tafel aus einer Urliste mit den abgerundet notierten Argumenten x' und mit der Gliederzahl $m = 10\,000$ entstanden; es gibt dann das Produkt $mH(x')$ jedesmal an, wie oft das betreffende x' in der Urliste vorkommt. Statt nun bei der Berechnung des Mittels $\Delta[f(x)]$ nach (1) jedes der m Glieder jener Formel einzeln anzusetzen, kann man die Glieder mit gleichem x' sofort zusammenziehen und demgemäß nach der Gleichung

$$\Delta[f(x)] = \sum f(x')H(x')$$

rechnen, wo die Summation über alle zwischen den Grenzen $\pm\infty$ enthaltenen Halbierungspunkte x' ausgedehnt werden darf, da ja für die „leeren" Tafelargumente die Zahlen $H(x')$ einfach null werden. Der vorstehende Ausdruck läßt sich noch weiter umformen. Bezeichnet man für den Augenblick mit y und z die beiden Wechselpunkte, die dem Tafelargument x' unmittelbar vorausgehen und nachfolgen, so ist

$$H(x') = \mathfrak{S}(z) - \mathfrak{S}(y) = \int_y^z \mathfrak{V}(x)\,dx,$$

woraus

$$f(x')\,H(x') = \int_y^z f(x')\,\mathfrak{V}(x)\,dx$$

folgt. Summiert man die vorstehende Gleichung über die ganze Verteilungstafel und beachtet, daß sich hierbei die Integrationsintervalle lückenlos aneinander schließen, so gelangt man zu dem Ausdrucke

$$\Delta[f(x)] = \int_{-\infty}^{\infty} f(x')\,\mathfrak{V}(x)\,dx, \tag{4}$$

worin x' den Halbierungspunkt derjenigen Teilstrecke bedeutet, in welcher sich die Integrationsveränderliche x bei ihrer Wanderung von $-\infty$ nach $+\infty$ jedesmal gerade befindet. Diesem Ausdrucke steht der von der Wirkung der Abrundung freie Durchschnittswert

$$\mathfrak{D}[f(x)] = \int_{-\infty}^{\infty} f(x)\mathfrak{V}(x)\,dx \tag{5}$$

gegenüber, so daß die Differenz (4) — (5) unmittelbar die Wirkung der Abrundung bei der Berechnung von $\mathfrak{D}(f)$ kennen lehrt.

Um bei der weiteren Rechnung nicht immer die Parameter c und h in den Formeln mitzuschleppen zu müssen, wollen wir im folgenden nur das in (3) angesetzte Hilfsargument u benutzen. Den Wechsel- und Halbierungspunkten von x und den zugehörigen Teilstreckenlängen entsprechen dann ebensolche durch (3) bestimmte Größen in dem Argument u. Setzt man ferner

$$f(x) = g(u), \qquad \mathfrak{V}(x)\,dx = V(u)\,du, \tag{6}$$

so liefert $V(u)$ die zu u gehörige Verteilungsfunktion und man erhält, wenn sich die Zeichen $\mathfrak{D}$ und Δ fortan auf $V(u)$ beziehen, statt (4) und (5) die Beziehungen

$$\mathfrak{D}[g(u)] = \int g(u)V(u)\,du, \qquad \Delta[g(u)] = \int g(u')V(u)\,du, \tag{7}$$

wo die Integrationen von $-\infty$ bis $+\infty$ laufen und u' den Halbierungspunkt derjenigen Teilstrecke bedeutet, in der sich die Integrationsveränderliche u jedesmal gerade befindet.

Bezeichnet man für das Argument u mit $2J$ die Teilstreckenlänge und mit a einen beliebig herausgegriffenen Halbierungspunkt, so lassen sich alle Halbierungspunkte von u in der Gestalt

$$u' = a + 2kJ \tag{8}$$

ansetzen, wo k alle positiven oder negativen ganzzahligen Werte anzunehmen hat. Die Größe a soll kurz als die *Phase* der Abrundung für das Argument u bezeichnet werden; ihr Wert kann offenbar auf das Intervall von 0 bis $2J$ oder auch auf das Intervall mit den Grenzen $\pm J$ beschränkt werden. Hiernach werden die Werte, die das Argument u innerhalb einer Teilstrecke mit dem Halbierungspunkt (8) durchläuft, durch den Ausdruck

$$u = a + 2kJ + z \tag{9}$$

dargestellt, wenn z von $-J$ nach $+J$ geht. Gleichzeitig sind die beiden den Halbierungspunkt (8) einschließenden Wechselpunkte durch

$$u = a + 2kJ - J \quad \text{und} \quad u = a + 2kJ + J$$

gegeben. Zerschneidet man nun das Integrationsgebiet der zweiten

Gleichung (7) in den Wechselpunkten und zerlegt dadurch das Integral in eine Summe von Teilintegralen, so erhalten wir die Darstellung

$$\Delta[g(u)] = \sum_k \int_{-J}^{J} g(a + 2kJ)\, V(a + 2kJ + z)\, dz, \tag{10}$$

die jetzt für verschiedene Formen der Funktion $g(u)$ näher zu untersuchen ist. Zuvor wollen wir uns jedoch eine unmittelbar anschauliche Vorstellung von der Bedeutung der Größe J zu verschaffen suchen.

§ 133. Die Größe $2J$ bedeutet die Teilstreckenlänge für das Hilfsargument u. Besitzt nun $2T$ dieselbe Bedeutung für das ursprüngliche Argument x, so folgt aus (3) die Beziehung $hT = J$. Wir wollen nun fortan die unschwer innezuhaltende Voraussetzung machen, daß der Parameter h sich nicht erheblich von seinem normalen Werte entferne. Dann ist die aus $1 = 2h^2s^2$ zu berechnende Größe s nahe gleich der Streuung des vorgelegten K.-G., und es wird

$$T = Js\sqrt{2}.$$

Führt man ferner neben J die für die folgenden Entwicklungen bequemere Größe P durch die Bedingung

$$2JP = \pi \tag{11}$$

ein, so wird

$$T = s\pi : P\sqrt{2},$$

woraus für den Quotienten $6s : 2T$ mit abgekürzten Zahlenwerten der Ausdruck

$$6s : 2T = 2.12 : J = 1.35P \tag{12}$$

folgt. Nun ist nach der in § 91 angestellten Erörterung bei den gewöhnlich vorkommenden K.-G. die Hauptmasse der möglichen Argumentwerte in einem Intervall enthalten, dessen Ausdehnung gleich der sechsfachen Streuung gesetzt werden darf. Hiernach gibt das Produkt $1.35P$ unmittelbar an, über wie viele Teilstrecken der vorgelegten Verteilungstafel sich jene Hauptmasse ausdehnt. Legt man an jeder Seite wegen der gewöhnlich überschießenden Extreme noch je eine Teilstrecke zu, so darf man sagen, daß die Zahl

$$N = 2 + 1.35P \tag{12.a}$$

näherungsweise angibt, wie viele Teilstrecken die Verteilungstafel für gewöhnlich vom ersten bis zum letzten vollen Argument umfassen wird. Ebenso läßt sich an der Hand eines Täfelchen, wie des folgenden

$N =$	5	6	7	8	9	10
$P =$	2.2	3.0	3.7	4.4	5.2	5.9
$N =$	11	12	13	14	15	16
$P =$	6.7	7.4	8.1	8.8	9.6	10.4

unmittelbar entnehmen, welchen Wert die Größe P ungefähr hat, wenn die Verteilungstafel N Teilstrecken umfaßt.

§ 134. Nach dieser Einschaltung kehren wir zu der Gleichung (10) zurück und setzen statt $g(u)$ das Polynom $\mathfrak{R}(u)_n$ ein, da es sich in erster Linie um die Auffindung der Größen

$$D_n = \mathfrak{D}[\mathfrak{R}(u)_n] \tag{13}$$

durch direkte Mittelbildung handelt. Statt des vorstehenden Sollwertes erhält man nun nach (10) durch die Operation Δ den Ausdruck

$$A(n) = \Delta[\mathfrak{R}(u)_n] = \sum_k \int_{-J}^{J} \mathfrak{R}(a + 2kJ)_n V(a + 2kJ + z)\, dz, \tag{14}$$

der offenbar von der Phase a, der Teilstreckenlänge $2J$ und den D-Koeffizienten der zu $V(u)$ gehörigen Φ-Reihe abhängt. Setzt man $a + 2J$ für a und gleichzeitig $k - 1$ für k, so bleibt die Gleichung (14) ungeändert. Folglich ist $A(n)$ eine periodische Funktion von a mit der Periode $2J$. Daraus ergibt sich nach der Theorie der *Fourier*schen Reihen für $A(n)$ eine trigonometrische Entwicklung von der Gestalt

$$A(n) = \sum_r A(n, r) \exp(2riaP), \tag{15.a}$$

in der für die Koeffizienten $A(n, r)$ die Integraldarstellung

$$2JA(n, r) = \int_{-J}^{J} A(n) \exp(-2riaP)\, da \tag{15.b}$$

besteht. Bei der Summation in (15.a) hat r von $-\infty$ bis $+\infty$ zu laufen, ferner sind die $A(n, r)$ für entgegengesetzte r konjugiert imaginär. Ersetzt man nun in (15.b) $A(n)$ durch den Ausdruck (14), so erhält man die Reihe

$$2JA(n, r) = \sum_k \int\int \mathfrak{R}(a + 2kJ)_n \exp(-2riaP) V(a + 2kJ + z)\, da\, dz.$$

Führt man $b = a + 2kJ$ als neue Integrationsveränderliche statt a ein, so wird unter Beachtung von (11)

$$2JA(n, r) = \sum_k \int\int \mathfrak{R}(b)_n \exp(-2ribP) V(b + z)\, db\, dz,$$

wo die Grenzen für die Integration nach b jetzt $(2k-1)J$ und $(2k+1)J$ werden. Da sich diese Grenzen für die einzelnen Glieder der nach k zu summierenden Reihe glatt aneinander schließen, während für z die Integrationsgrenzen beständig $-J$ und $+J$ sind, so läßt sich die Summe der Integrale in ein einziges Integral zusammenziehen, und man erhält

$$2JA(n, r) = \int_{-J}^{J} dz \int_{-\infty}^{\infty} \mathfrak{R}(b)_n \exp(-2ribP) V(b + z)\, db.$$

Führt man endlich noch $u = b + z$ als neue Veränderliche statt b ein, so wird

$$2JA(n, r) = \int_{-\infty}^{\infty} V(u)\,du \int_{-J}^{J} \mathfrak{R}(u - z)_n \exp(-2riuP + 2rizP)\,dz,$$

wofür wir auch den nach $V(u)$ zu nehmenden Durchschnitt

$$2JA(n, r) = \mathfrak{D}\Big[\int_{-J}^{J} \mathfrak{R}(u - z)_n \exp(-2riuP + 2rizP)\,dz\Big] \tag{16}$$

schreiben dürfen.

§ 135. Um jetzt die Integration nach z auszuführen, bilden wir für die $A(n, r)$ eine erzeugende Funktion $B(r)$ durch die Gleichung

$$B(r) = \sum_n A(n, r)(2v)^n, \qquad (n = 0, 1, 2, \ldots). \tag{17}$$

Ersetzt man hierin $A(n, r)$ durch den Ausdruck (16) und beachtet die Relation

$$\sum_n \mathfrak{R}(u - z)_n (2v)^n = \exp(-2uv + 2zv - v^2),$$

so wird

$$2JB(r) = \mathfrak{D}\Big[\int_{-J}^{J} \exp(-2uv + 2zv - 2riuP + 2rizP - v^2)\,dz\Big],$$

wo sich die Integration sofort ausführen läßt und mit den Abkürzungen

$$\exp(2vJ) - \exp(-2vJ) = 2JL, \qquad v + riP = U \tag{18}$$

den Ausdruck

$$B(r) = (-1)^r L(2U)^{-1} \mathfrak{D}[\exp(-2uU - v^2)] \tag{19}$$

liefert. Setzt man noch

$$F = \exp(-r^2P^2), \qquad M = \exp(2rivP), \tag{20}$$

so wird unter Berücksichtigung von (18)

$$B(r) = \sum_n A(n, r)(2v)^n = (-1)^r FLM(2U)^{-1} \mathfrak{D}[\exp(-2uU - U^2)]. \tag{21}$$

Setzt man in (14) für $V(u)$ die aus (2) und (6) folgende Reihe

$$2V(u) = \sum_n D_p \Phi(u)_{p+1}$$

ein, so stellt sich $A(n)$ als eine lineare Verbindung der D-Koeffizienten dar, und das Gleiche folgt aus (15.b) für die $A(n, r)$. Wir wollen demgemäß ansetzen

$$A(n, r) = \sum_p A(n, r, p) D_p, \qquad (p = 0, 1, 2, \ldots), \tag{22}$$

womit das zweite Glied von (21) in

$$\sum_p D_p \sum_n A(n, r, p)(2v)^n \tag{23}$$

übergeht. Andererseits hat man für die $\mathfrak{D}$-Größe des letzten Gliedes von (21) die Entwicklung

$$\mathfrak{D}[\exp(-2uU-U^2)] = \mathfrak{D}[\sum_p \mathfrak{R}(u)_p (2U)^p] = \sum_p D_p (2U)^p,$$

womit das genannte Glied die Gestalt

$$\sum_p (-1)^r FLM (2U)^{p-1} D_p$$

annimmt, aus der weiter durch Vergleichung mit (23) und Spaltung nach den D_p für die $A(n, r, p)$ die erzeugende Funktion

$$\sum_n A(n, r, p)(2v)^n = (-1)^r FLM(2U)^{p-1} \qquad (24)$$

folgt. Sind hierin die $A(n, r, p)$ durch Entwicklung der rechten Seite nach v gefunden, so erhält man schließlich für $A(n)$ die Darstellung

$$A(n) = \sum_r \sum_p A(n, r, p) D_p \exp(2riaP), \qquad (25)$$

die wir auch in der Gestalt

$$A(n) = \sum_p C(n, p) D_p, \qquad (26.a)$$

$$C(n, p) = \sum_r A(n, r, p) \exp(2riaP) \qquad (26.b)$$

schreiben können.

§ 136. Die Gleichung (25) enthält eine explizite Darstellung der $A(n)$ durch die bisher als bekannt angesehenen Größen, nämlich 1) die Koeffizienten D_p, 2) die Parameter c und h, 3) die für das System der Wechselpunkte maßgebenden Größen a und J oder P. Da der gefundene Zusammenhang zunächst etwas verwickelt ist, so wollen wir ihn etwas übersichtlicher gestalten, indem wir statt der verschiedenen A-Größen die Größen $E(n)$, $E(n, r)$ und $E(n, r, p)$ einführen und zwar durch die Gleichungen

$$2v \sum_n A(n)(2v)^n = L \sum_n E(n)(2v)^n, \qquad (27.a)$$

$$2v \sum_n A(n, r)(2v)^n = L \sum_n E(n, r)(2v)^n, \qquad (27.b)$$

$$2v \sum_n A(n, r, p)(2v)^n = L \sum_n E(n, r, p)(2v)^n, \qquad (27.c)$$

wo der in (18) eingeführte Faktor L nach Potenzen von v entwickelt zu denken ist. Man erhält dann, wenn man links und rechts nach v spaltet, für $A(n)$ den Ausdruck

$$A(n) = E(n) + \frac{1}{3!} J^2 E(n-2) + \frac{1}{5!} J^4 E(n-4) + \cdots, \qquad (28.a)$$

wo die Reihe abzubrechen ist, sobald man zu negativen Werten des Index von $E(n)$ gelangt. In der gleichen Weise wird

$$A(n, r) = E(n, r) + \frac{1}{3!} J^2 E(n-2, r) + \cdots, \qquad (28.b)$$

$$A(n, r, p) = E(n, r, p) + \frac{1}{3!} J^2 E(n-2, r, p) + \cdots. \qquad (28.c)$$

Hiermit verwandeln sich die aus (15) und (25) folgenden Entwicklungen

$$\sum_n A(n)(2v)^n = \sum_r \exp(2riaP) \sum_n A(n,r)(2v)^n \tag{29.a}$$

$$= \sum_p D_p \sum_r \exp(2riaP) \sum_n A(n,r,p)(2v)^n \tag{29.b}$$

in

$$\sum_n E(n)(2v)^n = \sum_r \exp(2riaP) \sum_n E(n,r)(2v)^n \tag{29.c}$$

$$= \sum_p D_p \sum_r \exp(2riaP) \sum_n E(n,r,p)(2v)^n, \tag{29.d}$$

aus denen

$$E(n) = \sum_r E(n,r) \exp(2riaP), \tag{30.a}$$

$$E(n,r) = \sum_p D_p E(n,r,p) \tag{30.b}$$

folgt. Hierfür läßt sich ähnlich wie in (26) auch schreiben

$$E(n) = \sum_p F(n,p) D_p, \tag{31.a}$$

$$F(n,p) = \sum_r E(n,r,p) \exp(2riaP). \tag{31.b}$$

Die E treten also einfach an die Stelle der A, jedoch nimmt die erzeugende Funktion der $E(n,r,p)$ nach (24) und (27.c) die Gestalt

$$\sum_n E(n,r,p)(2v)^n = (-1)^r (2v) FM(2U)^{p-1} \tag{32}$$

an, so daß der Faktor L in (24) jetzt durch den einfacheren Faktor $2v$ ersetzt worden ist. Hieraus sind nun die $E(n,r,p)$ herzuleiten, jedoch möge vorher eine Bemerkung eingeschaltet werden.

§ 137. Im allgemeinen hat bei den Bearbeitern beobachteter Kollektivreihen wohl niemals eine Unklarheit darüber bestanden, daß die Abrundung des Arguments bei der Aufstellung von Verteilungstafeln in die errechneten Resultate einen Fehler hineinbringt, sobald man das Verfahren der direkten Mittelbildung an den abgerundeten Argumenten ausführt; im besondern gibt *Fechner* über diesen Punkt eine ausführliche, in der Hauptsache auf numerischen Beispielen fußende Erörterung, die in seiner „Kollektivmaßlehre“ unter den Stichworten „Reduktionsstufe“ und „Reduktionslage“ zu finden ist. Meistens beruhigt man sich jedoch bei der Vorstellung, daß sich bei nicht allzu starker Zusammenziehung der Verteilungstafel das Zuviel und Zuwenig der Abrundung im großen und ganzen ausgleichen werde, und daß der verbleibende Fehlerrest schließlich nicht schädlicher sei, als der unvermeidliche Fehler, der bei beobachteten Reihen aus der unvollständigen Ausgleichung des Zufalls entspringt. Demgegenüber ist indessen zu bemerken, daß es doch zweierlei ist, ob ein Fehler aus unvermeidlichen und deshalb wohl oder übel zu ertragenden Ursachen ensteht, oder ob er darüber hinaus einen Bestandteil enthält, der erst

künstlich, d. h. durch den Rechnungsgang, hineingebracht worden ist. Es kann nicht zweifelhaft sein, daß man den letztgenannten der Sache fremden Bestandteil soweit auszumerzen hat, als das bei der Natur des Gegenstandes praktisch ausführbar ist. Damit werden wir zu der Aufgabe geführt, an der Hand der obigen Entwicklungen zu untersuchen, unter welchen Bedingungen man auf einem praktisch brauchbaren Wege von den durch die Abrundung gefälschten $A(n)$ zu den Sollwerten D_n gelangen könne.

§ 138. Bei der Behandlung der gestellten Frage wollen wir zunächst einmal voraussetzen, daß in den trigonometrischen Reihen, die in den obigen Entwicklungen auftreten, die periodischen Glieder unmerklich seien, daß man also höchstens die unperiodischen zum Index $r = 0$ gehörigen Glieder zu berücksichtigen habe. Dann folgt aus (29.d)

$$\sum_n E(n)(2v)^n = \sum_p D_p \sum_n E(n, 0, p)(2v)^n.$$

Andererseits folgt aus (32), weil für $r = 0$

$$F = M = 1, \qquad \dot{U} = v$$

wird, die Gleichung

$$\sum_n E(n, 0, p)(2v)^n = (2v)^p.$$

Die Verbindung der vorstehenden Gleichungen liefert

$$\sum_n E(n)(2v)^n = \sum_p D_p (2v)^p,$$

woraus sich

$$E(n) = D_n \tag{33}$$

und nach (28.a)

$$A(n) = D_n + \frac{1}{3!} J^2 D_{n-2} + \frac{1}{5!} J^4 D_{n-4} + \cdots \tag{34}$$

ergibt. Diese Beziehung gestattet ohne Schwierigkeit von den $A(n)$ auf die D_n überzugehen; sie enthält offenbar die Vorschrift zur Verbesserung der $A(n)$ wegen Abrundung, soweit nur die Teilstreckenlänge $2J$, aber nicht die Abrundungsphase a in Betracht kommt. Diese Bemerkung läßt sich noch etwas anders ausdrücken. In dem ursprünglichen Ausdrucke von $A(n)$ liefern die rein periodischen Glieder einen Bestandteil, der bei wechselnden Phasenwerten um Null herumschwankt. Daher stellt die aus (34) folgende Verbesserung von $A(n)$ einen Mittelwert aus den für wechselnde Phasen geltenden vollständigen Werten der Verbesserung dar.

§ 139. Wollen wir nun weiter die Verbesserung wegen der Phase erhalten, so haben wir auf die Gleichungen (32) zurückzugehen. Führt man für die Binomialkoeffizienten die Bezeichnung

$$(p-1)_f = (p-1)! : [f!(p-1-f)!]$$

ein, so hat man zur Entwicklung der rechten Seite von (32) die beiden Reihen

$$(2U)^{p-1} = (2v + 2riP)^{p-1} = \sum_f (p-1)_f (2v)^f (2riP)^{p-1-f},$$
$$M = \exp(2rivP) = \sum_k [(riP)^k (2v)^k : k!]$$

auszumultiplizieren, dann den Faktor $(-1)^r 2vF$ hinzuzufügen und schließlich zur Ermittlung von $E(n, r, p)$ die Glieder mit $(2v)^n$ herauszusuchen. Dies liefert einen Ausdruck, den wir aus folgenden Gleichungen zusammensetzen wollen:

$$E(n, r, p) = i^{2r+p+n+2} G(n, r, p) K(n, r, p), \tag{35.a}$$
$$G(n, r, p) = 2^{p-1} F(rP)^{p+n-2}, \tag{35.b}$$
$$K(n, r, p) = \sum_f U_f, \qquad (f = 0, 1, 2, \ldots), \tag{35.c}$$
$$U_f = (p-1)_f (-Q)^f : (n-1-f)!, \tag{35.d}$$
$$Q = 1 : 2r^2 P^2. \tag{35.e}$$

Die Summation nach f in (35.c) bricht, wie man sich leicht überzeugt, für $p = 1, 2, \ldots$ bei der kleineren der beiden Zahlen $p - 1$ und $n - 1$ von selber ab, geht dagegen bis $n - 1$ für $p = 0$.

Hiernach besitzen die vollständigen Ausdrücke der Koeffizienten $F(n, p)$ in den Gleichungen (31), vermittelst deren die D_p aus den $E(n)$ zu ermitteln sein würden, keineswegs eine besonders einfache Gestalt. Es wäre nun trotzdem immerhin denkbar, wenn auch nicht wahrscheinlich, daß die Gleichungen (31) eine glatte und für die Rechnung hinlänglich bequeme analytische Auflösung zuließen, indessen steht eine solche, falls sie überhaupt vorhanden ist, im Augenblicke nicht zur Verfügung, und man ist deshalb, wenn man die $A(n)$ von einem merklichen Einflusse auch der periodischen Glieder befreien will, darauf angewiesen, die Koeffizienten $F(n, p)$ zu berechnen und dann die Auflösung nach den D-Größen rein numerisch durchzuführen. Das wäre jedoch in Wirklichkeit ein außerordentlich weiter Umweg. Denn wenn man einmal zu der numerischen Auflösung linearer Gleichungen greifen will, so läßt sich das erstrebte Ziel weit einfacher dadurch erreichen, daß man unmittelbar mit der früher aufgestellten Gleichung (2) operiert. Denn dort werden die gesuchten D-Koeffizienten in einen unmittelbaren Zusammenhang mit den Zahlen der Verteilungstafel gebracht, ohne daß man erst nötig hätte, die Größen $A(n)$, $E(n)$ und $F(n, p)$ zu berechnen. Außerdem enthalten die für alle Wechselpunkte angesetzten Gleichungen (2) alles, was sich über die D_p auf Grund der vorgelegten Verteilungstafel aussagen läßt: mehr kann man aber auch nicht aus den Werten der $A(n$ herausholen.

Bei dieser Sachlage hat man, falls man bei dem Verfahren der direkten Mittelbildung bleiben will, zwischen zwei Dingen die Wahl: entweder läßt man die periodischen Glieder ganz unberücksichtigt, auf die Gefahr hin, unter Umstäuden merkliche Fehler zu begehen, oder aber man sucht das Ding so einzurichten, daß diese lästigen Glieder in der Tat vernachlässigt werden dürfen. Will man letzteren Fall herbeiführen, so hat man die Frage zu beantworten, wann die $E(n, r, p)$ für ein von Null verschiedenes r unmerklich werden. Betrachtet man nun die in (35) gegebenen Ausdrücke, so läßt sich zwar der Verlauf des Faktors $G(n, r, p)$ unschwer übersehen, dagegen gestaltet sich der Verlauf der Größen $K(n, r, p)$ um so mannigfaltiger, je größer n wird. Hält man, um zu bestimmteren Vorstellungen zu gelangen, zunächst die Indizes n, r, p fest und läßt P beständig wachsen, so gehen die E-Größen, wegen des Verhaltens des Exponentialfaktors F, sehr rasch gegen Null. Damit ist aber nichts gewonnen, denn die Frage dreht sich gerade um die Beträge, die bei kleinen oder doch mittleren Werten von P zum Vorschein kommen. Läßt man ferner bei festen n, p, P den Index r wachsen, so gehen die E wiederum rasch gegen Null, d. h. die in unserer Entwicklung vorkommenden trigonometrischen Reihen konvergieren sehr rasch. Damit ist aber wiederum nichts gewonnen, weil die lästige Verwicklung der Gleichungen (31) schon mit den niedrigsten, zu $r = 1$ gehörigen Gliedern eintritt. Hält man drittens r, p, P fest, während n wächst, so gehen die E gegen Null, und zwar ähnlich wie der Quotient $(rP)^n : n!$. Auch dieser Umstand ist ohne Bedeutung, da ein merklicher Einfluß der periodischen Glieder in den niedrigeren, $A(n)$ gerade so unbequem ist wie in den höheren. Läßt man endlich bei festen n, r, P den Index p wachsen, so gehen die E schließlich über alle Grenzen ins Unendliche. Allerdings wird dieses Anwachsen dadurch kompensiert, daß die $E(n, r, p)$ mit den gegen Null konvergierenden D_p zu multiplizieren sind; gleichzeitig erkennt man aber auch, daß sich garnicht ohne weiteres übersehen läßt, wie diese beiden Umstände gegeneinander wirken und daß darum das Verhalten der periodischen Glieder ganz wesentlich von der Beschaffenheit des vorgelegten Verteilungsgesetzes ahhängen wird.

§ 140. Wenn nun nach der letzten Bemerkung eine zugleich *allgemeine* und *bestimmte* Aussage nicht zu erlangen ist, so ist man darauf angewiesen, durch die Untersuchung besonderer Fälle Anhaltspunkte für die Beantwortung der gestellten Frage zu gewinnen.

Da die Unsicherheit über das Verhalten der periodischen Glieder wesentlich davon herrührt, daß die höheren D_p stark vergrößert in jene Glieder eingehen, so wird man den für die Unmerklichkeit der periodischen Bestandteile günstigsten Fall erhalten, wenn die D_p mit Ausnahme von D_0 sämtlich verschwinden, wenn also $V(u)$ in das

einfache E.-G. übergeht. Nun verwandelt sich in (35.d) für $p = 0$ der Binomialkoeffizient $(p-1)^f$ in $(-1)^f$, so daß die Reihe $K(n, r, 0)$ nur positive Glieder enthält. Berücksichtigen wir der Einfachheit halber in den $K(n, r, 0)$ nur das erste Glied, nämlich $1 : (n-1)!$, so erhalten wir damit sicher keine zu großen Werte von $K(n, r, 0)$. Beachtet man ferner, daß bei den Sinus oder Kosinus der Winkel $2raP$ die Größen $E(n, r, 0)$ und $E(n, -r, 0)$ gleichzeitig ihren Beitrag zu den Koeffizienten liefern, so kommt es offenbar darauf an, die Größenordnung des Ausdruckes

$$2G(n, r, 0) : (n-1)! = \exp(-r^2P^2)(rP)^{n-2} : (n-1)! \qquad (36)$$

zu ermitteln. Zu dem Ende enthält das nachstehende unmittelbar verständliche Täfelchen für etliche Wertepaare von n und rP die mit ihren Vorzeichen angesetzten gemeinen Logarithmen des Ausdruckes (36).

$rP =$	2	3	4	5	6	7
$n = 1$	— 2.0	— 4.4	— 7.6	— 11.6	— 16.4	— 22.1
2	— 1.7	— 3.9	— 6.9	— 10.9	— 15.6	— 21.3
3	— 1.7	— 3.7	— 6.6	— 10.4	— 15.2	— 20.7
4	— 1.9	— 3.7	— 6.5	— 10.2	— 14.6	— 20.4
5	— 2.2	— 3.6	— 6.5	— 10.1	— 14.7	— 20.1
6	— 2.6	— 4.1	— 6.6	— 10.1	— 14.6	— 20.0
7	— 3.1	— 4.4	— 6.8	— 10.2	— 14.6	— 19.9
8	— 3.6	— 4.7	— 7.0	— 10.4	— 14.7	— 19.9
9	— 4.2	— 5.2	— 7.3	— 10.6	— 14.8	— 20.0
10	— 4.9	— 5.7	— 7.7	— 10.8	— 15.0	— 20.1

Die vorstehenden Zahlen gestatten nun einen Schluß auf den kleinsten zulässigen Wert von P. Wenn nämlich die periodischen Glieder unmerklich sind, so fallen, wie wir oben gesehen haben, die $E(n)$ mit den D_n zusammen; es müssen also, da im vorliegenden Falle die D_n null sind, die $E(n)$ und ebenso die $2E(n, r, 0)$ unmerklich sein. Setzt man nun mit Rücksicht auf die in der Kollektivmaßlehre gewöhnlich innegehaltene Rechnungsschärfe fest, daß die Merklichkeitsgrenze in den $E(n)$ durch die Einheit der vierten Stelle, also durch 0.0001 gegeben sei, so sind die ungünstigen Werte von P aus der Bedingung zu bestimmen, daß für sie die oben angesetzten Logarithmen algebraisch größer als — 4.0 ausfallen. Das findet für die niedrigsten periodischen Glieder, also für $r = 1$, zwischen $P = 3$ und $P = 4$ statt, und zwar nahe bei $P = 3$. Dazu gehört aber nach dem in § 133 gegebenen Täfelchen die Teilstreckenanzahl 6 bis 7. Hieraus ziehen wir den Schluß, daß bei dem einfachen E.-G. und bei der angenommenen Rechnungsschärfe die periodischen Glieder merklich werden, sobald die Verteilungstafel vom ersten bis zum letzten „vollen" Argument weniger als 8 Teilstrecken umfaßt. Diese Zahl bezeichnet zugleich

die äußerste überhaupt vorkommende Grenze, da wir ja für $V(u)$ die günstigste Form gewählt haben, und da ferner die obigen Tafelwerte wegen der mit den $K(n, r, 0)$ vorgenommenen Abkürzung etwas zu klein ausgefallen sind.

§ 141. Wir wollen nun weiter einen Fall betrachten, der erheblich ungünstiger liegt, ohne jedoch gerade ein Extrem darzustellen. Es möge der oberhalb der Abszissenachse liegende Verlauf der Verteilungskurve ein zum Nullpunkte symmetrisches und gleichschenkliges Dreieck bilden, mit der Basis $2g$ und der Höhe gk. Danach ist außerhalb der Abszissengrenzen $\pm g$ die Funktion $V(u)$ gleich Null, innerhalb dagegen

$$V(u) = k[g - u \operatorname{sg}(u)]. \tag{37}$$

Bestimmt man g und k aus den Bedingungen $3k = 1$ und $g^2 = 3$, so wird die Dreiecksfläche, wie es sein muß, gleich Eins. Ferner wird $\mathfrak{D}(u) = 0$ und $\mathfrak{D}(u^2)$, wie man sich leicht überzeugt, gleich 0.5. Daraus folgt für den Parameter h der Normalform der Wert $h = 1$, so daß das Hilfsargument der Φ-Funktionen in der Normalform mit u zusammenfällt. Um nun die $E(n, r)$ zu bilden, greifen wir auf die Gleichungen (17) und (19) zurück und haben zunächst

$$\sum_n A(n, r)(2v)^n = (-1)^r L(2U)^{-1}\mathfrak{D}[\exp(-2uU - v^2)],$$

woraus durch Übergang auf die E-Größen

$$\sum_n E(n, r)(2v)^n = (-1)^r(2v)(2U)^{-1}\exp(-v^2)\mathfrak{D}[\exp(-2uU)] \tag{38}$$

folgt. Hierin bilden wir nun den Durchschnitt nach der Gleichung

$$\mathfrak{D}[\exp(-2uU)] = \int_{-g}^{g} \exp(-2uU) k[g - u \operatorname{sg}(u)]\, du.$$

Die Ausführung der Integration und die Einsetzung in (38) liefert mit der Abkürzung

$$W = \exp(2gU - v^2) + \exp(-2gU - v^2) - 2\exp(-v^2)$$

die Darstellung

$$\sum_n E(n, r)(2v)^{2n} = (-1)^r(2kv)(2U)^{-3}W. \tag{39}$$

Die Größe W läßt sich ohne Schwierigkeit entwickeln, ebenso die Potenz von $2U$. Ich gehe jedoch nicht darauf ein, weil es für unseren Zweck schon genügt, den Ausdruck für $E(1, r)$ anzusetzen. Man erhält nach einer kleinen Reduktion

$$2E(1, r) = i^{2r-1}\sin(grP)^2 : 3(rP)^3,$$

worin man, um für $r = 1$ die Genauigkeitsgrenze 0.0001 innezuhalten, für P den Wert 14.9 zu setzen hätte. Daraus ergäbe sich nach § 133 (12.a), daß die Verteilungstafel über wenigstens 22 Teilstrecken

ausgedehnt sein müßte. Wollte man als Grenze nur die halbe Einheit der dritten Stelle vorschreiben, so käme man auf $P = 11.9$ und auf die Teilstreckenanzahl 18.

Das letzte Beispiel ist absichtlich so gewählt worden, daß die Umstände für die Abminderung der periodischen Glieder schon ziemlich ungünstig liegen. Das ist dadurch erreicht worden, daß die Tangente der Verteilungskurve $V(u)$ ihre Richtung mehrmals unstetig ändert. Infolgedessen gehen in der Φ-Reihe für $V(u)$ die D-Koeffizienten, die man wegen $E(n, 0) = D_n$ unschwer aus (39) findet, für größere n verhältnismäßig langsam gegen Null, so daß man von der Reihe eine ziemlich große Anzahl von Gliedern mitzunehmen hat, bevor der Rest gänzlich unmerklich wird. Trotzdem ist der für die Teilstreckenanzahl gefundene Grenzwert noch keineswegs unbequem groß: wir werden später die Rechnung für die direkte Mittelbildung in eine Gestalt bringen, bei der selbst 30 oder 40 Teilstrecken noch keine erhebliche Erschwerung der Arbeit bedeuten. Da bei den gewöhnlich vorkommenden stetigen Kollektivreihen die Umstände meistens günstiger liegen als bei dem behandelten Beispiele, so wird es zulässig sein, als *Faustregel* für die innezuhaltende Teilstreckenanzahl das Mittel aus den gefundenen Grenzen 8 und 22, also die Zahl 15 anzusetzen, mit dem Vorbehalt, in zweifelhaften Fällen darüber hinauszugehen. Wie sich der Rechner helfen kann, wenn ihm die Verteilungstafel in stark zusammengezogener Gestalt vorgelegt wird, soll später bei den numerischen Beispielen auseinandergesetzt werden.

§ 142. Die vorstehende Entwicklung lehrt, daß es im allgemeinen keine Schwierigkeit bietet, die periodischen Glieder durch passende Ansetzung der Wechselpunkte unschädlich zu machen. In diesem Falle sind dann nur noch die unperiodischen Bestandteile in den $A(n)$ nach (34) zu beseitigen. Zu dem Ende gehen wir auf die Gleichung (27.a) zurück, in der jetzt wegen der Unmerklichkeit der periodischen Glieder D_n für $E(n)$ geschrieben werden darf, so daß man die Gleichung

$$2v \sum_n A(n)(2v)^n = L \sum_n D_n (2v)^n \tag{40}$$

erhält, wo nach (18) die Größe L durch die Gleichung

$$2JL = \exp(2Jv) - \exp(-2Jv)$$

bestimmt ist. Hiernach erhält man die D-Reihe, wenn man die A-Reihe mit der Entwicklung

$$2v : L = 1 - a_1(2Jv)^2 + a_2(2Jv)^4 - a_3(2Jv)^6 + a_4(2Jv)^8 - \cdots \tag{41}$$

multipliziert, wo

$$a_1 = 1 : 6, \qquad a_2 = 7 : 360, \tag{42.a}$$

$$a_3 = 31 : 15120, \qquad a_4 = 127 : 604800, \tag{42.b}$$

$$\cdots \qquad \cdots$$

ist. Damit wird

$$D_n = A(n) - a_1 J^2 A(n-2) + a_2 J^4 A(n-4) - a_3 J^6 A(n-6) + \cdots. \quad (43)$$

Statt übrigens die $A(n)$ auf die D_n zu reduzieren, kann man auch die Potenzmittel von $x - c$ wegen Abrundung verbessern. Da nämlich

$$A(n) = \Delta[\mathfrak{R}(u)_n], \qquad D_n = \mathfrak{D}[\mathfrak{R}(u)_n] \quad (43.a)$$

ist, so läßt sich (40) in der Gestalt

$$2v\Delta[\sum_n \mathfrak{R}(u)_n (2v)^n] = L\mathfrak{D}[\sum_n \mathfrak{R}(u)_n (2v)^n]$$

oder

$$2v\Delta[\exp(-2uv - v^2)] = L\mathfrak{D}[\exp(-2uv - v^2)]$$

schreiben, woraus

$$2v\Delta[\exp(-2uv)] = L\mathfrak{D}[\exp(-2uv)] \quad (43.b)$$

folgt. Setzt man nun zur Abkürzung

$$n!P(n) = \Delta[(x-c)^n], \qquad n!Q(n) = \mathfrak{D}[(x-c)^n], \quad (44)$$

so erhält man wegen $u = h(x-c)$ und $hT = J$ die Beziehungen

$$2v\sum_n (-1)^n P(n)(2hv)^n = L\sum_n (-1)^n Q(n)(2hv)^n,$$

$$2v : L = 1 - a_1(2Thv)^2 + \cdots.$$

Die Formeln (40) und (43) bleiben also bestehen, wenn man darin die Größen $A(n)$, D_n, v und J mit den Größen $(-1)^n P(n)$, $(-1)^n Q(n)$, hv und T vertauscht. Im besonderen wird

$$Q(0) = P(0) = 1. \quad (45.a)$$

$$Q(1) = P(1), \quad (45.b)$$

$$Q(2) = P(2) - \tfrac{1}{6}T^2 P(0). \quad (45.c)$$

Aus der zweiten dieser Gleichungen folgt $\Delta(x-c) = \mathfrak{D}(x-c)$ oder

$$\Delta(x) = \mathfrak{D}(x), \quad (46)$$

d. h. *der Argumentdurchschnitt wird durch die direkte Mittelbildung richtig gefunden, sobald die periodischen Glieder unmerklich sind.* Nimmt man mit Rücksicht hierauf für c seinen Normalwert $\mathfrak{D}(x)$ und bezeichnet die nach den Operationen Δ und $\mathfrak{D}$ berechneten Streuungen mit S und s, so folgt aus (45.c)

$$s^2 = S^2 - \frac{1}{3}T^2. \quad (47)$$

Die Streuung S ist also stets größer als ihr Sollwert s und die erforderliche Reduktion ist bei den gewöhnlichen vorkommenden Abrundungen garnicht so unbedeutend. Setzt man z. B. $2T = 1$ und $s = 3$, so geht die sechsfache Streuung über 18 Teilstrecken und mindestens ebensoviele Strecken treten in der Verteilungstafel auf, so daß die Abrundung keineswegs groß ist. Gleichwohl erhält man

$S = 3.014$, also eine Abweichung von dem Sollwerte, die bei der gewöhnlich innegehaltenen Rechnungsschärfe durchaus zu berücksichtigen sein würde.

§ 143. Der hier erörterte Zusammenhang zwischen den Größen $A(n)$ und D_n bezieht sich ausschließlich auf die stetigen K.-G., denn bei den unstetigen K.-G. sind die Operationen Δ und $\mathfrak{D}$ identisch, wie sich unmittelbar aus der Definition des $\mathfrak{D}$-Zeichens ergibt. Darin liegt ein wesentlicher Unterschied zwischen den beiden Arten von Kollektivreihen begründet, den wir bei den früheren Untersuchungen der Einheitlichkeit halber absichtlich immer in den Hintergrund geschoben haben. Denkt man sich eine Summentafel mit äquidistanten Argumentwerten und Wechselpunkten gegeben, so kann diese, wenn nichts weiter hinzugefügt wird, ebensowohl einem stetigen, wie einem unstetigen K.-G. angehören. Im Falle der Unstetigkeit liefern die durch direkte Mittelbildung gefundenen Größen $A(n)$ sogleich die D-Koeffizienten der Φ-Reihe für die Summenfunktion $\mathfrak{S}(x)$, während in dem anderen Falle an die $A(n)$ erst noch die Reduktion auf die D_n anzubringen ist, die sich, wie wir gesehen haben, bei genügend kleiner Abrundung ohne Schwierigkeit ermitteln läßt. Man erhält also zwei verschiedene Φ-Reihen, je nachdem die vorgelegte Tafel als die Summentafel eines stetigen oder unstetigen K.-G. behandelt wird. Der Unterschied zwischen den beiden Reihen wird jedoch nach den obigen Entwicklungen um so geringer, je größer die Menge der vollen Tafelargumente ist.

Da die Summenkurve für einen unstetigen K.-G. die Gestalt einer Treppe besitzt, und da sich ferner die Summe der Produkte $A(n)\,\Phi(u)_n$, je mehr Glieder der Reihe mitgenommen werden, um so inniger dem gebrochenen Zuge der Treppenlinie anschmiegen muß, so wird die Reihe im allgemeinen nur langsam konvergieren, selbst wenn die niedrigeren Koeffizienten rasch auf geringe Beträge herabgehen, denn mit einer mäßigen Anzahl von Reihengliedern ist wegen des Verhaltens der Funktionen Φ_n jener treppenförmige Verlauf nur roh zu erreichen. Bei den stetigen K.-G. fällt der angegebene Grund für die langsame Konvergenz offenbar fort, so daß es, soweit der Anschluß der Reihe an die Zahlenwerte der vorgelegten Summentafel in Betracht kommt, im allgemeinen vorteilhaft sein wird, die unstetigen K.-G. auf das Schema der stetigen zu bringen.

Bei einem unstetigen K.-G. ist der Verlauf der Summentreppe durch die beobachteten $\mathfrak{U}(x)$ vollständig bestimmt. Infolgedessen ist auch das ganze System der numerischen Elemente völlig bestimmt, und man erhält die Zahlenwerte der Elemente unmittelbar durch die Berechnung der Größen $A(n)$. Bei einem stetigen K.-G. dagegen liefert die Beobachtung zunächst nichts weiter als eine gewisse Anzahl diskreter Punkte der Summenkurve, und es bleibt für die Verbindung

dieser Punkte durch eine stetige Kurve ein größerer oder geringerer Spielraum übrig. Daraus folgt, daß auch das System der Elemente vorläufig nicht völlig bestimmt sein kann. Will man nun trotzdem ein bestimmtes Wertsystem erhalten, so muß die Unbestimmtheit durch Hinzufügung weiterer Festsetzungen irgendwie aufgehoben werden. In dem hier vorliegenden Falle geschieht das, wie man erkennt, dadurch, daß man zuerst das völlig bestimmte System der $A(n)$ ableitet und darauf unter Vernachlässigung der Phasenwirkung die ebenfalls völlig bestimmten Verbesserungen anbringt, die nur von der Länge der Teilstrecken abhängen.

§ 144. Die Ergebnisse der bisherigen Untersnchung geben noch zu einer anderen Bemerkung Anlaß. Bei der Mitteilung beobachteter stetiger Kollektivreihen wird aus Unkenntnis der für die Sicherheit der Rechnung wesentlichen Erfordernisse sehr oft dadurch gesündigt, daß man die Verteilungstafel, um ihr ein „besseres" Aussehen zu geben, stark zusammenzieht und so das Ergebnis einer manchmal recht mühsamen Arbeit geradezu illusorisch macht. Denkt man sich aus der Verteilungstafel die Summentafel abgeleitet und für jeden Wechselpunkt die Gleichung

$$2\mathfrak{S}(x) - 1 = \sum_n D_n \Phi(u)_n$$

angesetzt, so enthält das hierdurch entstehende Gleichungssytem alles, was sich auf Grund der beobachteten Zahlen über die numerischen Elemente des K.-G. aussagen läßt. Enthält nun, wie das vorkommt, das System nur zwei oder drei für die numerische Rechnung brauchbare Gleichungen, so tappt man bezüglich der Werte der numerischen Elemente einfach im dunkeln, denn man kann dann die Normalwerte von c und h innerhalb eines gewissen Spielraums ganz beliebig annehmen und trotzdem durch passende Wahl der niedrigsten D-Koeffizienten die gegebenen Ordinaten der Summenkurve genau darstellen. Noch deutlicher wird der begangene Fehler, wenn man die Aufgabe geometrisch formuliert. Die Summenkurve, deren Bildungsgesetz im allgemeinen nicht näher bekannt ist, soll ihrem Verlaufe nach auf Grund der durch die Beobachtung gegebenen Punkte festgelegt werden: es ist also ein *Fehler*, dem Rechner nur zwei oder drei Punkte mitzuteilen, wenn die Beobachtung mehr Punkte geliefert hat oder bei sachgemäßer Anordnung liefern konnte. Wir werden später auf diesen Gegenstand nochmals zurückzukommen haben.

Siebenzehnte Vorlesung.

Das gewöhnliche Urnenschema und seine Erweiterung.

§ 145. Der Inhalt der letzten Vorlesungen bezog sich auf Fragen allgemeiner Natur, bei denen es auf die besondere Beschaffenheit der betrachteten K.-G. nicht weiter ankam. Im folgenden wollen wir uns nun mit einigen besonderen Kollektivreihen beschäftigen; den Beginn soll hierbei die Untersuchung von gewissen theoretisch konstruierten Verteilungen machen. Derartige Reihen treten jedesmal auf, wenn man feststellen will, ob eine beobachtete Verteilung sich diesem oder jenem Urnenschema anschließen läßt, denn man hat dann die beobachteten Zahlen mit derjenigen Verteilung zu vergleichen, welche aus dem angenommenen Urnenschema entspringt.

Bei der Behandlung der gestellten Aufgaben wird sich zeigen, daß die in den früheren Abschnitten entwickelte Darstellung von Verteilungen durch die Φ-Reihe eine natürliche Ausgestaltung der Methode bildet, die von *Laplace* für die Untersuchung von Funktionen großer Zahlen angegeben worden ist.

Als erstes Beispiel soll folgende Aufgabe dienen. Aus einer Urne, die weiße und schwarze Kugeln enthält, wird n-mal unter Zurücklegung der Kugel gezogen und als beobachtetes x die rH. der gezogenen weißen Kugeln notiert, wobei x offenbar einen der $n + 1$ Werte

$$0 : n, \quad 1 : n, \quad \ldots \quad n : n$$

annehmen kann. Diesen aus n „Zügen" bestehenden „Versuch" denken wir uns im ganzen m-mal angestellt und erhalten so aus den notierten x einen unstetigen K.-G. vom Umfange m mit einer gewissen Verteilung $\mathfrak{U}(x)$ und dem Argument x. Denkt man sich weiter m unendlich groß, so gehen die $\mathfrak{U}(x)$, falls der Satz von der gleichmäßigen Erschöpfung der gleichmöglichen Fälle zutrifft, in die Größen $\mathfrak{W}(x)$ über, wo $\mathfrak{W}(x)$ entsprechend der früher eingeführten Bezeichnungsweise die $\mathfrak{W}$. bedeutet, mit der x bei dem einzelnen Versuche zu erwarten ist. Nun lassen sich aber die $\mathfrak{W}(x)$ unmittelbar aus den Bedingungen des Urnenschemas berechnen, so daß man aus der Vergleichung der theoretischen $\mathfrak{W}(x)$ mit den bei einem endlichen m beobachteten $\mathfrak{U}(x)$ erkennen kann, wieweit sich bei der ausgeführten Versuchsreihe die Wirkungen des Zufalls ausgeglichen haben. Sind ferner die Werte $\mathfrak{U}(x)$ nicht aus einem wirklichen Urnenversuche, sondern aus irgend einer anderen statistisch beobachteten Massenerscheinung entstanden, so lehrt die Vergleichung von $\mathfrak{U}(x)$ und $\mathfrak{W}(x)$, ob und wieweit jene Massenerscheinung auf das hier betrachtete Urnenschema reduziert werden darf. Damit gelangt man dann auf

den Weg, den *Lexis* eingeschlagen hat, um die in älterer Zeit häufig vorgekommenen kritiklosen Anwendungen der W.-R. auf das richtige Maß zu bringen.

Bei der gestellten Aufgabe führen die einzelnen Werte von $\mathfrak{W}(x)$ nach der in § 18 gegebenen Entwicklung auf die Glieder der Binomialreihe für $(p+q)^n$, wo p die $\mathfrak{W}$. für den einzelnen weißen Zug bedeutet, und q gleich $1-p$ ist. Diese Darstellung der $\mathfrak{W}(x)$ ist natürlich für kleine Werte von n zugleich die einfachste und bequemste, wird dagegen für ein größeres n undurchsichtig, so daß eine Umformung zweckmäßig wird, zu der uns die Φ-Reihe führen wird. Zuvor wollen wir jedoch die gestellte Aufgabe etwas verallgemeinern. Statt nämlich innerhalb eines Versuches aus einer einzigen Urne n-mal zu ziehen, kann man auch, ohne an den $\mathfrak{W}(x)$ etwas zu ändern, aus n Urnen gleicher Füllung je einmal ziehen. In diesem Falle kann man aber dann noch einen Schritt weiter gehen und annehmen, daß die n Urnen verschiedene Füllungen enthalten. Damit gelangen wir zu folgender Aufgabe. Gegeben sind n Urnen; die Füllung der Urne mit der Nummer k ist durch die für den einzelnen Zug geltenden Größen

$$\mathfrak{W}(\text{weiß}) = p_k, \quad \mathfrak{W}(\text{schwarz}) = q_k, \quad p_k + q_k = 1 \tag{1}$$

bestimmt; der einzelne Versuch umfaßt jedemal n Züge, und zwar je einen aus jeder Urne; bei jedem Versuch wird als beobachtetes x die rH. der gezogenen weißen Kugeln notiert: gesucht wird das theoretische Verteilungsgesetz $\mathfrak{W}(x)$ für das Argument x.

§ 146. Betrachten wir zunächst innerhalb einer Reihe von Versuchen nur die Züge aus der k-ten Urne und bezeichnen mit x_k die bei einem Zuge auftretende rH. der weißen Kugeln, so ist x_k teils gleich Null, teils gleich Eins. Bedeutet ferner $\mathfrak{U}(x_k)$ die zu x_k gehörige theoretische Verteilung, so ist für die beiden möglichen Argumentwerte

$$\mathfrak{U}(0) = q_k, \quad \mathfrak{U}(1) = p_k,$$

woraus für ein beliebiges c die auf $\mathfrak{U}$ bezügliche Durchschnittsgleichung

$$\mathfrak{D}[(x_k - c)^r] = \mathfrak{U}(0)(0-c)^r + \mathfrak{U}(1)(1-c)^r$$

folgt. Mit $c = p_k$ wird daher für $r = 1$ und $r = 2$

$$\mathfrak{D}(x_k - p_k) = 0, \quad \mathfrak{D}[(x_k - p_k)^2] = p_k q_k,$$

und weiter

$$\mathfrak{D}(x_k) = p_k, \quad \operatorname{str}(x_k)^2 = p_k q_k. \tag{2}$$

Betrachten wir nun weiter die n Züge eines Versuchs gleichzeitig, so ist nach den getroffenen Festsetzungen

$$nx = x_1 + x_2 + \cdots + x_n. \tag{3}$$

Da hierin die x_k innerhalb der Reihe der Versuche unabhängig voneinander nach ihren Verteilungen $\mathfrak{U}(x_k)$ variieren, so ist nx ein aus den Argumenten x_k gemischtes Argument, und wir können unmittelbar die in der XIII. Vorlesung gefundenen Sätze anwenden. Hierbei werde zur Abkürzung

$$np = \sum_k p_k, \quad nq = \sum_k q_k, \quad p + q = 1, \tag{4}$$

$$r_k = p_k - p, \quad \sum_k r_k = 0, \quad \sum_k r_k^2 = nP_2 \tag{5}$$

eingeführt, und das $\mathfrak{D}$-Zeichen, wie es in jener Vorlesung geschah, jedesmal auf die Verteilung der dahinter stehenden Argumente bezogen. Dann folgen nach den Formeln (17) und (19) in § 103 aus (3) und (2) die Beziehungen

$$\mathfrak{D}(nx) = \sum_k \mathfrak{D}(x_k) = \sum_k p_k = np,$$

$$\operatorname{str}(nx)^2 = \sum_k \operatorname{str}(x_k)^2 = \sum_k p_k q_k$$

$$= \sum_k (p + r_k)(q - r_k) = npq - \sum_k r_k^2,$$

und weiter

$$\mathfrak{D}(x) = p, \quad n \operatorname{str}(x)^2 = pq - P_2. \tag{6}$$

Die Größe P_2 ist offenbar das Streuungsquadrat der p_k, wenn diese als Glieder eines K.-G. aufgefaßt werden.

Da die Argumentmischung bei wachsendem n dem einfachen Exponentialgesetz zustrebt, sobald die Beträge der $\operatorname{str}(x_k)$ merklich von derselben Größenordnung sind, so erhält man in diesem Falle für die zu $\mathfrak{W}(x)$ gehörende Summenfunktion $\mathfrak{S}(x)$ die genäherte Darstellung

$$2\mathfrak{S}(x) - 1 = \Phi[h(x - p)], \tag{7}$$

wo h aus der in (6) gegebenen Streuung nach

$$2h^2(pq - P_2) = n \tag{8}$$

zu berechnen ist. Die Formel (7) enthält bereits den sogenannten *Poisson*schen Satz, dessen Vervollständigung offenbar die Hinzufügung der von den Koeffizienten $D_3, D_4, \ldots$ abhängenden Glieder erfordern würde. Diese Glieder können ebenfalls nach dem Verfahren der XIII. Vorlesung gefunden werden, ergeben sich jedoch auf etwas bequemere Weise, wenn wir uns von vornherein die besondere Gestalt von $\mathfrak{W}(x)$ zu nutze machen.

§ 147. Bildet man mit der willkürlichen Größe u und den vorhin eingeführten Urnenkonstanten p_k, q_k das Produkt

$$P = (up_1 + q_1)(up_2 + q_2) \cdots (up_n + q_n) \tag{9}$$

und entwickelt nach u in die Reihe

$$P = \sum_a P(a) u^a, \qquad (a = 0, 1, \cdots n)$$

so ist nach den in § 18 entwickelten Sätzen $P(a)$ die $\mathfrak{W}$. dafür, daß bei einem Versuche a weiße Kugeln gezogen werden. Hiernach ist $P(a)$ identisch mit $\mathfrak{W}(x)$, sobald $a = nx$ gesetzt wird. Man hat also auch

$$P = \sum_x \mathfrak{W}(x) u^{nx}, \qquad (nx = 0, 1, \cdots n).$$

In der vorstehenden Summe sind die einzelnen Werte der von x abhängenden Funktion u^{nx} mit den entsprechenden Verteilungszahlen $\mathfrak{W}(x)$ multipliziert, folglich ist diese Summe gleich dem nach $\mathfrak{W}(x)$ genommenen Durchschnitt aus u^{nx}, so daß, wenn das $\mathfrak{D}$-Zeichen sich fortan auf $\mathfrak{W}(x)$ bezieht

$$\mathfrak{D}(u^{nx}) = P = (up_1 + q_1) \cdots (up_n + q_n) \tag{10}$$

wird.

Setzt man $u = 1 - t$, so wird mit Rücksicht auf die in (4) und (5) eingeführten Bezeichnungen

$$\log (up_k + q_k) = \log (1 - tp_k) = \log [1 - t(p + r_k)]. \tag{11}$$

Sieht man nun die Größe

$$L(0) = \log (1 - tp) \tag{12}$$

als Funktion von p an und bezeichnet ihre sukzessiven Ableitungen nach p mit $L(1)$, $L(2)$, ..., so folgt aus (11) nach dem *Taylor*schen Lehrsatze die Reihenentwicklung

$$\log (up_k + q_k) = \sum_g \frac{1}{g!} L(g) r_k^g, \quad (g = 0, 1, \ldots). \tag{13}$$

Setzt man ferner

$$\log P = n Q, \qquad \sum_k r_k^a = n P_a, \tag{14}$$

wo offenbar $P_0 = 1$ und $P_1 = 0$ ist, so liefert die Summation von (13) nach k die Gleichung

$$Q = \sum_g \frac{1}{g!} L(g) P_g. \tag{15}$$

Die $L(g)$ erscheinen, wenn man $L(0)$ auf Grund der Gleichung (12) nach Potenzen von tp entwickelt und dann g-mal nach p differentiiert, als Potenzreihen von t und p. Führt man ferner statt u und t die neue Veränderliche v durch

$$u = \exp v, \quad t = 1 - \exp v \tag{16}$$

ein, so läßt sich $L(g)$ als eine Potenzreihe von p und v darstellen, die in der Gestalt

$$L(g) = g! \sum_k A(k, g)\, v^k \tag{17}$$

geschrieben werden möge, wo die $A(k, g)$ Potenzreihen von p sind. Die Einsetzung in (15) liefert die Darstellung

$$Q = \sum_g \sum_k A(k, g)\, v^k P_g, \tag{18}$$

in der die Summation nach k erst mit $k = 1$ zu beginnen braucht, da für $v = 0$ die Werte $t = 0$, $u = 1$ und wegen (9) und (14) die Werte $P = 1$, $Q = 0$ eintreten. Aus (18) fließt dann, wenn man nach v ordnet und

$$A(k) = A(k, 0) P_0 + A(k, 1) P_1 + \cdots \tag{19.a}$$

setzt, die Darstellung

$$Q = A(1) v + A(2) v^2 + \cdots. \tag{19.b}$$

Die Gleichung (17) nimmt für $g = 0$ die Gestalt

$$L(0) = \sum_k A(k, 0) v^k$$

an. Differentiiert man diese Beziehung g-mal nach p, so erhält man links $L(g)$, folglich muß man rechter Hand auf die in (17) angesetzte Reihe kommen. Demnach ist $g!A(k, g)$ die g-te Ableitung von $0!A(k, 0)$ nach p, so daß es ausreicht, zunächst die Größen $A(k, 0)$ aufzusuchen. Zu dem Ende bilden wir unter Berücksichtigung von (12), (16) nnd (17) die Beziehung

$$\log(1 - p + p \exp v) = L(0) = \sum_k A(k, 0) v^k,$$

aus der durch partielle Differentiation nach p und v die Gleichungen

$$[\exp v - 1] : [q + p \exp v] = \frac{\partial L}{\partial p} = \sum_k \frac{\partial A(k, 0)}{\partial p} v^k,$$

$$p \exp v : [q + p \exp v] = \frac{\partial L}{\partial v} = \sum_k k A(k, 0) v^{k-1}$$

folgen. Multipliziert man die beiden vorstehenden Gleichungen mit pq und -1 und summiert, so geben die Bestandteile links den Wert $-p$, während auf der rechten Seite der Ausdruck

$$\sum_k pq \frac{\partial A(k, 0)}{\partial p} v^k - \sum_k k A(k, 0) v^{k-1}$$

entsteht, wo k in beiden Summen mit $k = 1$ zu beginnen hat. Spaltet man nun links und rechts nach v, so erhält man zunächst aus den von v freien Gliedern die Gleichung

$$p = A(1, 0). \tag{20}$$

Die übrigen Glieder führen zu der Rekursionsformel

$$(k + 1) A(k + 1, 0) = (p - p^2) \frac{\partial A(k, 0)}{\partial p}, \tag{21}$$

in der k die Werte $1, 2, \ldots$ anzunehmen hat.

Die Berechnung der niedrigeren $A(k, 0)$ läßt sich nach (21) rasch ausführen und liefert zunächst Polynome von p. Differentiiert man dann $A(k, 0)$ g-mal nach p und dividiert darauf durch $g!$, so erhält man $A(k, g)$. Daraus ergeben sich dann nach (19.a) die $A(k)$. Setzt man zur Abkürzung

$$r = q - p = 1 - 2p = 2q - 1, \quad t = pq,$$

so ergeben sich, da $P_0 = 1$ und $P_1 = 0$ ist, bis zum Index 6 hin die Ausdrücke

$$1!A(1) = p, \tag{22.a}$$

$$2!A(2) = t - P_2, \tag{22.b}$$

$$3!A(3) = rt - 3rP_2 + 2P_3, \tag{22.c}$$

$$4!A(4) = t - 6t^2 - (7 - 36t)P_2 + 12rP_3 - 6P_4, \tag{22.d}$$

$$\begin{aligned} 5!A(5) = {} & r(t - 12t^2) - r(15 - 120t)P_2 \\ & + (50 - 240t)P_3 - 60rP_4 + 24P_5, \end{aligned} \tag{22.e}$$

$$\begin{aligned} 6!A(6) = {} & t - 30t^2 + 120t^3 - (31 - 540t + 1800t^2)P_2 \\ & + r(180 - 1200t)P_3 - (390 - 1800t)P_4 + 360rP_5 - 120P_6. \end{aligned} \tag{22.f}$$

Aus den vorstehenden Gleichungen ist die allgemeine Gestalt der $A(k)$ bereits zu erkennen, soll jedoch nicht weiter verfolgt werden.

§ 148. Faßt man die Formeln (16), (10), (14) und (19.b) zusammen, so kann man schreiben

$$\mathfrak{D}[\exp(nxv)] = P, \quad \log P = nQ,$$
$$Q = A(1)v + A(2)v^2 + \cdots,$$

wo die $A(k)$ bis zum Index 6 aus (22) zu entnehmen sind. Setzt man ferner

$$T = P \exp[-nA(1)v - nA(2)v^2],$$

so ist einerseits

$$T = \mathfrak{D}[\exp(nxv - nA(1)v - nA(2)v^2)], \tag{23.a}$$

andrerseits

$$\log T = nQ - nA(1)v - nA(2)v^2 = nA(3)v^3 + nA(4)v^4 + \cdots. \tag{23.b}$$

Geht man in der letzten Gleichung vom Logarithmus zum Numerus über, so erhält man für T eine nach Potenzen von n und v fortschreitende Reihe, die wir in der Gestalt

$$\begin{aligned} T = 1 & + n\,[(3, 1)v^3 + (4, 1)v^4 + \cdots] \\ & + n^2\,[(6, 2)v^6 + (7, 2)v^7 + \cdots] \\ & + n^3\,[(9, 3)v^9 + (10, 3)v^{10} + \cdots] + \cdots \end{aligned}$$

oder kürzer

$$T = 1 + \sum_g \sum_k (g, k)\, v^g n^k, \qquad (g, k = 1, 2, 3 \cdots) \tag{24}$$

schreiben wollen, wobei zu beachten ist, daß die (g, k) verschwinden, sobald g kleiner als $3k$ ist. Im besondern ist

$$(3, 1) = A(3), \quad (4, 1) = A(4), \quad (5, 1) = A(5), \tag{25.a}$$

$$(6, 2) = \tfrac{1}{2} A(3)^2, \quad (7, 2) = A(3)A(4), \tag{25.b}$$

$$(9, 3) = \tfrac{1}{6} A(3)^3. \tag{25.c}$$

Führt man jetzt die Parameter c, h und die Veränderliche w durch die Gleichungen

$$c = p, \quad 2h^2(pq - P_2) = n, \quad 2wh = -nv \tag{26}$$

ein, so läßt sich das Argument der Exponentialfunktion in (23.a) unter Berücksichtigung der aus (22) folgenden Werte von $A(1)$ und $A(2)$ in der Gestalt

$$-2h(x - c)w - w^2$$

schreiben. Damit erhält man aus (23.a)

$$T = \mathfrak{D}[\exp(-2h(x - c)w - w^2)] = \sum_g \mathfrak{D}[\mathfrak{R}(hx - hc)_g](2w)^g$$

oder, wenn D_g die Koeffizienten in der mit den Parametern c, h gebildeten Φ-Reihe für die betrachtete Verteilung $\mathfrak{W}(x)$ bezeichnet,

$$T = \sum_g D_g (2w)^g.$$

Hält man hiergegen die aus (24) und (26) fließende Gleichung

$$T = 1 + \sum_g \sum_k (g, k)(-h)^g n^{k-g}(2w)^g,$$

so ergibt sich

$$n^g D_g = (-h)^g \sum_k (g, k) n^k, \qquad (k = 1, 2, \cdots) \tag{27}$$

Da die Größen $(1, k)$ und $(2, k)$ null sind, so verschwinden D_1 und D_2, d. h. die angesetzten Parameterwerte gehören der Normalform an, und man hat in Übereinstimmung mit (6)

$$c = \mathfrak{D}(x) = p, \quad 2h^2 \operatorname{str}(x)^2 = 1, \tag{28.a}$$

$$n \operatorname{str}(x)^2 = pq - P_2. \tag{28.b}$$

Um das Ergebnis noch etwas übersichtlicher zu machen werde gesetzt

$$nN^2 = 1, \quad hN = H, \quad 2H^2(pq - P_2) = 1, \tag{29}$$

wo H offenbar für ein unbegrenzt wachsendes n endlich bleibt. Damit nimmt (27) die Gestalt

$$D_g = (-HN)^g \sum_k (g, k) n^k$$

an. Multipliziert man die vorstehende Gleichung mit $\Phi(u)_g$, wo u das Hilfsargument $h(x - c)$ bedeutet, und summiert darauf nach g

von $g = 3$ an, so entsteht die Φ-Reihe für die betrachtete Verteilung $\mathfrak{W}(x)$, jedoch ohne das Anfangsglied $\Phi(u)$. Demnach ist

$$2\mathfrak{S}(x) - 1 - \Phi(u) = \sum\sum(g, k)(-HN)^g n^k \Phi(u)_g.$$

Da hierin die Summation nach g erst mit $g = 3k$ zu beginnen braucht, so setzen wir $3k + l$ statt g, wo l von Null an zu laufen hat, und gewinnen die Reihe

$$\sum_k\sum_l(3k + l, k)(-H)^{3k+l}\Phi_{3k+l}N^{k+l}$$

oder, wenn wir $f = k + l$ setzen,

$$\sum_f\sum_k(f + 2k, k)(-H)^{f+2k}\Phi_{f+2k}N^f,$$

wo f von Eins an läuft. Daraus folgt für die Summenfunktion der untersuchten Verteilung die Darstellung

$$2\mathfrak{S}(x) - 1 = \Phi(u) + \sum_f\sum_k(f + 2k, k)(-H)^{f+2k}\Phi(u)_{f+2k}N^f,$$

die sich, wenn man nach N ordnet, aus den Gleichungen

$$2\mathfrak{S}(x) - 1 = \Phi(u) + (1)N + (2)N^2 + (3)N^3 + \cdots, \tag{30.a}$$

$$(f) = \sum_k(f + 2k, k)(-H)^{f+2k}\Phi(u)_{f+2k} \tag{30.b}$$

zusammensetzen läßt. Es wird also z. B. für die niedrigsten Glieder

$$\left.\begin{aligned}(1) &= (3, 1)(-H)^3\Phi_3,\\ (2) &= (4, 1)(-H)^4\Phi_4 + (6, 2)(-H)^6\Phi_6,\\ (3) &= (5, 1)(-H)^5\Phi_5 + (7, 2)(-H)^7\Phi_7 + (9, 3)(-H)^9\Phi_9,\end{aligned}\right\} \tag{31}$$

wo die angesetzten (g, k) aus (25) und (22) zu entnehmen sind.

§ 149. Die Gleichungen (30) enthalten die Lösung der gestellten Aufgabe und lassen die Gestalt der Glieder erkennen, die man zur Vervollständigung der in (7) gegebenen abgekürzten Formel nötig hat. Bei der numerischen Anwendung von (30) ist zu beachten, daß in $\mathfrak{W}(x)$ unter x die möglichen Argumentwerte von x, also die Vielfachen von $1 : n$ zu verstehen sind. Demgemäß wird man nach einer früheren Bemerkung aus $\mathfrak{S}(x)$ die Summentafel für die Halbierungspunkte zwischen den Argumentwerten von $\mathfrak{W}(x)$ herleiten und daraus durch Subtraktion der benachbarten Summengrößen die Verteilungstafel, d. h. die Tafel der $\mathfrak{W}(x)$ bilden.

Die Urnenzahl n war bei der vorstehenden Entwicklung keiner besonderen Einschränkung unterworfen worden, so daß man für n auch z. B. den Wert Eins annehmen darf. Allerdings konvergiert für kleine Werte von n die Reihe entsprechend langsam; ebenso findet eine Verlangsamung der Konvergenz statt, wenn H große Werte annimmt, was z. B. sicher eintritt, sobald eine der beiden Zahlen p, q sehr klein ausfällt. In der Regel braucht jedoch die Zahl n

keineswegs besonders groß zu sein, damit eine ausreichend rasche Abnahme der Reihenglieder stattfinde. Das ist deswegen von Bedeutung, weil ausgedehnte Versuchsreihen mit großem n nicht immer leicht zu beschaffen sind. Andrerseits ist das Glied mit Φ_3 für die Urnenzahl $n = 1000$, die schon als ziemlich groß gelten darf, nicht immer unmerklich. Nehmen wir der Einfachheit halber an, daß die Urnenkonstanten p_k zusammenfallen, und daß demgemäß die Potenzenmittel $P_2, P_3, \ldots$ verschwinden, so besitzt das Glied $D_3 \Phi_3$ die Gestalt

$$-\tfrac{1}{6} Npq(q-p) H^3 \Phi_3 = -N(q-p)\Phi_3 : 12\sqrt{2pq}.$$

Dies liefert z. B. für $p = 0.1$ und $q = 0.9$ den Ausdruck $-0.157 N\Phi_3$, woraus in $\mathfrak{S}(x)$ für das Maximum von Φ_3 das Glied $+0.177N$ entspringt. Das gibt für $n = 1000$ den Wert 0.0056, der nicht immer zu vernachlässigen sein wird.

Ist n so groß, daß die Φ-Reihe auf ihr Anfangsglied beschränkt werden darf, so beruht die Berechnung von $\mathfrak{W}(x)$, um alles übersichtlich zusammen zu stellen, auf den Gleichungen

$$c = \mathfrak{D}(x) = p, \quad n \operatorname{str}(x)^2 = pq - P_2, \quad 2h^2 \operatorname{str}(x)^2 = 1, \tag{32.a}$$

$$u = h(x-c), \quad 2\mathfrak{S}(x) - 1 = \Phi(u). \tag{32.b}$$

Diese Formeln enthalten die sogenannte *Poisson*sche Verallgemeinerung des Satzes von *Bernoulli*. Der *Bernoulli*sche Satz selber ergibt sich, wenn die Urnenkonstanten p_k untereinander und mit p zusammenfallen, wobei die Größen $P_2, P_3, \ldots$ verschwinden. Es wird dann

$$c = \mathfrak{D}(x) = p, \quad n \operatorname{str}(x)^2 = pq, \quad 2h^2 \operatorname{str}(x)^2 = 1, \tag{33.a}$$

$$u = h(x-c), \quad 2\mathfrak{S}(x) - 1 = \Phi(u). \tag{33.b}$$

Wir wollen nun auf die Bedeutung dieser Formeln noch etwas näher eingehen.

Gewöhnlich erscheint der *Bernoulli-Poisson*sche Satz in den Darstellungen der W.-R. als eine Art Krönung des Gebäudes, d. h. als ein Abschluß der allgemeinen theoretischen Entwicklungen, und wird dabei mit einem gewissen Überschwang als das *Gesetz der großen Zahlen* bezeichnet. Die oben gegebene Herleitung rückt den Satz in die richtige Beleuchtung: man erkennt, daß es sich um einen besonderen und verhältnismäßig einfachen Fall einer allgemeinen Beziehung handelt, die über das untersuchte Urnenschema weit hinausreicht.

Beschränken wir uns zunächst auf die Betrachtung des *Bernoulli*schen Urnenschemas, so lassen sich an der zugehörigen Verteilungsfunktion folgende Merkmale hervorheben: 1) die Annäherung an das einfache Exponentialgesetz bei wachsendem n, 2) die Abnahme der Streuung bei wachsendem n, 3) der Umstand, daß die numerischen Elemente von $\mathfrak{W}(x)$ schon durch das erste unter ihnen, nämlich $p = \mathfrak{D}(x)$, voll-

ständig bestimmt sind. Die beiden ersten Merkmale werden gewöhnlich als besonders wichtig betont; es wird sich jedoch zeigen, daß das dritte als die Hauptsache anzusehen ist. Gehen wir nun weiter die aufgeführten Punkte im einzelnen durch, so läßt sich folgendes bemerken.

§ 150. Die Annäherung an das E.-G. hat ihre Quelle in der Argumentmischung. Wenn man also das immerhin etwas anspruchsvoll klingende Wort „Gesetz" gebrauchen will, so hätte der Satz über das Verhalten der Argumentmischungen sicherlich mehr Anspruch auf eine solche Bezeichnung, als die *Bernoulli*sche Formel. Aber selbst wenn man sich auf die Betrachtung des Urnenschemas beschränken wollte, würde der Name „Gesetz" dem vollständigen Ausdruck der Funktion $\mathfrak{W}(x)$ und nicht dem Anfangsgliede der Φ-Reihe zuzusprechen sein, denn die Beschränkung auf das Anfangsglied bei hinreichend großem n ist in Wahrheit doch nur eine Annehmlichkeit für den Rechner, die sich dieser selbstverständlich zunutze machen wird, wenn die Umstände danach angetan sind. Damit büßt offenbar das Beiwort „groß" in dem *Bernoulli*schen Satze erheblich an Bedeutung ein.

Das zweite Merkmal, nämlich die Abhängigkeit der Streuung von n, entspringt ebenfalls aus den Eigenschaften der Argumentmischung und hat geschichtlich eine wichtige, wenn auch nicht gerade glückliche Rolle gespielt. Bedeutet a eine beliebig kleine aber angebbare positive Konstante, so ist die $\mathfrak{W}$., daß bei einem Versuche das beobachtete x zwischen die Grenzen $p \pm a$ falle, durch die Differenz

$$\mathfrak{S}(p+a) - \mathfrak{S}(p-a) = \Phi(ha)$$

gegeben. Diese Differenz geht mit unbegrenzt wachsendem n gegen Eins, weil ja in diesem Falle die Größe h ebenfalls über alle Grenzen wächst. Da man nun sagen kann, daß sich innerhalb eines Versuches der Zufall um so vollständiger ausgeglichen habe, je näher das beobachtete x an p liegt, so erhält man den Satz, daß bei dem betrachteten Urnenschema die „mathematische Wahrscheinlichkeit" für die Ausgleichung des Zufalls mit wachsendem n gegen Eins konvergiere. Daraus wurde nun unvermerkt der Satz, daß eine solche Ausgleichung bei beobachteten großen Zahlen mit Notwendigkeit eintrete. Hierbei übersah man, daß zwei Dinge scharf zu trennen sind, nämlich einerseits die von der Erfahrung unabhängige Geltung der aus dem Begriffe der $\mathfrak{W}$-Größen deduktiv gewonnenen Lehrsätze einer rein mathematischen Häufigkeitsrechnung, andrerseits die nur auf induktivem Wege zu rechtfertigende Anwendung jener Lehrsätze auf beobachtete Vorgänge.

Übler jedoch als die falsche Deutung eines an sich einwandfreien Lehrsatzes war der Umstand, daß man, weil in der *Bernoullischen* Formel von großen Zahlen die Rede ist, aus dieser Formel ein rich-

tiges Prokrustesbett machte, in das sich die Reihen der Statistik, sobald sie nur recht große Zahlen enthielten, unweigerlich mußten einspannen lassen. Der unvermeidliche Rückschlag gegen ein derartiges unkritisches Verfahren ist denn auch nicht ausgeblieben.

§ 151. Die vorhin an dritter Stelle aufgeführte Eigenschaft, daß bei dem *Bernoullischen* Urnenschema die Funktion $\mathfrak{W}(x)$ nur einen einzigen unabhängigen Parameter enthält, ist erst ziemlich spät in der richtigen Weise ausgenutzt worden. Als Erläuterung hierzu kann der in § 57 erwähnte Versuch von *R. Wolf* über das Nadelproblem dienen. Der Versuch bestand darin, daß eine dünne Nadel 5000 mal blindlings auf ein Gitter von äquidistanten Parallelen geworfen wurde, wobei der einzelne Wurf als weißer oder schwarzer Zug galt, je nachdem die Nadel eine Gittergerade traf oder nicht traf. Aus den gewählten Abmessungen von Nadel und Gitter folgte zunächst der theoretische Wert $p = 0.5093$, woraus nach der *Bernoullischen* Formel mit $n = 5000$ die Streuung gleich 0.0071 gefunden wird. Die Beobachtung ergab für x den Wert 0.5064, so daß die Abweichung von dem Sollwerte 0.0029 beträgt und erheblich kleiner als die Streuung ist; man darf also sagen, daß Beobachtung und Rechnung gut miteinander stimmen, und daß eine weitgehende Ausgleichung des Zufalls stattgefunden hat. Ein solches Ergebnis ist immerhin von Interesse, gibt aber, da nur ein einziger Versuch vorliegt, keine Auskunft über die durchaus nicht überflüssige Vorfrage, wie weit denn eigentlich die einzelnen Nadelwürfe den Zügen des *Bernoullischen* Schema gleichgestellt werden dürfen. Wenn ferner, wie das bei den Beobachtungsreihen der Statistik der Fall ist, der Wert von p nicht im voraus bekannt ist, so führt offenbar der einzelne Versuch nicht wesentlich über die Angabe des beobachteten x hinaus. Man muß also, um mehr zu erfahren, die Beobachtungen anders anlegen oder doch anders behandeln, als dies bei den vorhin als Beispiel angeführten Zahlen geschehen ist. Der Erste, der in dieser Beziehung auf dem Gebiete der eigentlichen Statistik planmäßig den richtigen Weg verfolgt hat, ist, wie früher erwähnt wurde, *Lexis* gewesen. Die Grundgedanken seines Verfahrens sind aus der Arbeit zu entnehmen, die wir schon einmal (§ 65 flg.) benutzt haben. Hier genügt es, eines der behandelten Beispiele in unserer Ausdrucksweise zu skizzieren, wobei es unwesentlich ist, daß bei *Lexis* der Gang der Rechnung im einzelnen anders verläuft, als im folgenden vorausgesetzt ist.

Die Geburtenmengen, die während eines gewissen Zeitraums innerhalb fester Zeitstrecken, z. B. von Monat zu Monat, und innerhalb fester Bezirke registriert worden sind, mögen als Glieder eines K.-G. angesehen und entsprechend zusammenfaßt werden; ferner werde für jedes Glied dieses K.-G. als Argument x die rH. einer Knabengeburt notiert. Die Umordnung der Urliste liefert dann eine gewisse

beobachtete Verteilung $\mathfrak{U}(x)$, zu der man nach irgend einem zulässigen Verfahren, sagen wir durch direkte Mittelbildung, die „beobachteten" numerischen Elemente $\mathfrak{D}(x)$, $\operatorname{str}(x)$, D_3, D_4, ... aufsucht. Setzt man nun voraus, daß die einzelne Geburt auf das Schema der Einzelziehung aus einer Urne konstanter Füllung reduziert werden dürfe, und läßt man ferner die Schwankungen in der jedem beobachteten x zugrunde liegenden Zügezahl n beiseite, so ist dem beobachteten $\mathfrak{U}(x)$ eine theoretische Verteilung $\mathfrak{W}(x)$ zuzuordnen, die aus den oben entwickelten Formeln zu entnehmen ist und nur einen einzigen unabhängigen Parameter, nämlich die mit p bezeichnete $\mathfrak{W}$. einer Knabengeburt, enthält. Ist die gemachte Voraussetzung zulässig, so muß $\mathfrak{U}(x)$ mit $\mathfrak{W}(x)$ bis auf die Reste unausgeglichener Zufälligkeiten übereinstimmen, sobald in $\mathfrak{W}(x)$ für p der passende numerische Wert eingesetzt wird. Nimmt man nun für p das beobachtete $\mathfrak{D}(x)$ und berechnet daraus die übrigen Elemente von $\mathfrak{W}(x)$, so liefert die Vergleichung der beobachteten und der theoretischen Elemente $\operatorname{str}(x)$, D_3, D_4, ... je ein Kriterium für die Zulässigkeit des *Bernoullischen* Urnenschemas. Stimmt die Vergleichung bei allen Elementen in befriedigender Weise, so ist das Schema zulässig; dagegen ist das Schema als unbrauchbar abzulehnen, sobald auch nur bei einem einzigen der Elemente ein unzulässig großer Widerspruch zwischen Beobachtung und Rechnung auftritt.

Statt die Elemente $\operatorname{str}(x)$, D_3, D_4, ... einzeln zu vergleichen, kann man auch, was in der Regel anschaulicher ist, ihre Gesamtwirkung ins Auge fassen, indem man, wenn die Streuungen stimmen, die zu $\mathfrak{U}(x)$ und $\mathfrak{W}(x)$ gehörigen Verteilungs- und Summenkurven miteinander vergleicht.

Das vorstehend skizzierte Beispiel läßt erkennen, daß es für die Sicherheit der Untersuchung wichtig ist, die Menge der Versuche zu steigern, d. h. den Umfang der für $\mathfrak{U}(x)$ benutzten Kollektivreihe möglichst groß zu machen. Dagegen kommt es keineswegs darauf an, die Zügezahl n, d. h. die Menge der für das einzelne x benutzten Einzelfälle besonders groß zu wählen. Es genügt vielmehr, n nur so groß zu nehmen, daß der Rechner zur Erleichterung der Arbeit sich auf das erste Glied der Φ-Reihe beschränken darf. Oder kürzer ausgedrückt: *die Menge der Versuche ist wichtiger als ihre Länge.*

Weiter erkennt man nun aber auch, worin der Kern des von *Lexis* befolgten Verfahrens besteht. *An die Stelle der früher als selbstverständlich betrachteten Annahme, daß man die mit Zufälligkeiten behafteten Massenerscheinungen der Statistik ohne weiteres auf ein Urnenschema reduzieren dürfe, tritt jetzt die Aufgabe, an der Hand bestimmter Kriterien die Zulässigkeit jener Annahme jedesmal besonders zu prüfen.*

Die von *Lexis* geforderte Prüfung setzt voraus, daß die nötigen Kriterien jedesmal auch wirklich aufgestellt werden können. Hierbei läßt sich nun erkennen, welche Rolle bei dem *Bernoullischen* Schema der Umstand spielt, daß das zugehörige $\mathfrak{W}(x)$ nur einen unabhängigen Parameter enthält. Angenommen, das mit dem beobachteten $\mathfrak{U}(x)$ zu vergleichende $\mathfrak{W}(x)$ enthielte zwei Parameter — sagen wir $\mathfrak{D}(x)$ und str(x) — die unabhängig voneinander variieren können, dann würde bei der Vergleichung von $\mathfrak{U}(x)$ und $\mathfrak{W}(x)$ das Kriterium fortfallen, das in dem vorhin betrachteten Beispiel durch die Streuung geliefert wird. In der gleichen Weise würde jeder weitere unabhängige Parameter die Anzahl der ursprünglich vorhandenen Kriterien um eine Einheit vermindern, bis man schließlich an den Punkt gelangt, wo die Untersuchung über die Zulassung oder Ablehnung des Urnenschemas mit einem Fragezeichen endigt, da ja die Anzahl der aus $\mathfrak{U}(x)$ mit einiger Sicherheit abzuleitenden Elemente praktisch stets eine begrenzte ist.

Bei der eingehenderen Untersuchung über das Geschlechtsverhältnis der einfachen Geburten gelangt *Lexis* zu dem Schlusse, daß sich die Beobachtungen in befriedigender Weise dem *Bernoullischen* Schema fügen. Ein solches Ergebnis besitzt offenbar, sobald es durch vielfältige Proben den Rang einer allgemeinen Erfahrung erlangt hat, eine Bedeutung, die über den Bereich der eigentlichen Statistik hinaus in das Gebiet anderer Wissenschaften, z. B. der Physiologie, übergreift, denn die physiologische Erklärung der betrachteten Massenerscheinung würde das Resultat der statistischen Beobachtung mit zu berücksichtigen haben. Das führt uns zu der Frage, wie weit in einem solchen Falle die Annahme einer merklich konstanten $\mathfrak{W}$. für die einzelnen weißen und schwarzen Züge nach Ausweis der Beobachtung nicht nur erlaubt, sondern auch geboten sei. Denn man hat sich gegenwärtig zu halten, daß die Übereinstimmung zwischen einer beobachteten und einer zugeordneten theoretischen Verteilung schon dann als vorhanden angesehen wird, wenn die übrig bleibenden unvermeidlichen Abweichungen nach Größe und Gruppierung aus der unvollständigen Ausgleichung des Zufalls erklärt werden dürfen. Es ist also recht wohl denkbar, daß zwei sehr verschiedene Ansätze des Urnenschemas gleich gut mit einer beobachteten Verteilung stimmen, weil die beiden theoretischen Verteilungen nur geringe Abweichungen voneinander besitzen. Bei der Untersuchung der gestellten Frage wollen wir zunächst auf den in den letzten Betrachtungen beiseite gelassenen *Poissonschen* Fall näher eingehen.

§ 152. Bei dem *Poissonschen* Ansatz ist die $\mathfrak{W}$. des einzelnen weißen Zuges nicht konstant, sondern durchläuft innerhalb des n Züge umfassenden Versuches die Werte $p_1, p_2, \ldots p_n$. Dieses Wertsystem kehrt von Versuch zu Versuch wieder, wobei es allerdings nicht not-

wendig ist, daß die p_k stets in derselben Reihenfolge auftreten. Die Streuung ergibt sich, wie wir gesehen haben, aus den Gleichungen

$$n\,\mathrm{str}\,(x)^2 = pq - P_2, \tag{34.a}$$

$$np = \sum_k p_k, \quad q = 1 - p, \quad nP_2 = \sum_k (p_k - p)^2, \tag{34.b}$$

wobei P_2 als das Streuungsquadrat der zu einem K.-G. vereinigten p_k angesehen werden darf. Betrachtet man nun den nach dem *Bernoullischen* Ansatz aus

$$n\,\mathrm{str}\,(x)^2 = pq$$

berechneten Betrag der Streuung als den normalen Wert, so wird in dem *Poissonschen* Falle die Streuung notwendig unternormal, weil P_2 wesentlich positiv ist.

Es ist bemerkenswert, daß die wiederkehrende Schwankung in der W. der weißen Züge auf die Streuung dieselbe Wirkung ausübt, die aus einer Verbundenheit der betrachteten Massenerscheinung entspringen würde. Man denke sich z. B., daß in dem *Bernoullischen* Falle die Ziehungen nicht völlig blindlings erfolgten, und daß die ziehende Person bis zu einem gewissen Grade die Möglichkeit besäße, bei jedem Zuge auf das Ergebnis $x = p$ hinzuwirken, dann würde dadurch die Abweichung der beobachteten x von p durchschnittlich vermindert, also die Streuung auf einen unternormalen Wert gebracht werden.

Das in (34.a) auftretende Glied P_2 kann bewirken, daß $\mathrm{str}\,(x)$ beliebig klein und sogar null wird. Wenn z. B. die p_k, was ja nach den Voraussetzungen des Schemas nicht ausgeschlossen ist, teils null, teils Eins sind, wenn also jede Urne nur Kugeln von einer Farbe enthält, so liefert jeder Versuch denselben Wert von x, so daß $\mathrm{str}\,(x)$ in der Tat verschwindet.

Der *Poissonsche* Ansatz, der sich experimentell unschwer verwirklichen läßt, kann auch bei natürlichen Massenerscheinungen recht wohl vorkommen. Man denke sich z. B., daß die einzelnen Züge mit konstantem zeitlichen Intervall aufeinander folgen und jahresweise in Versuche zusammengefaßt werden; ferner sei die W. des einzelnen weißen Zuges von der Jahreszeit — sagen wir von der Lufttemperatur — abhängig, dann wird bei den Urnenkonstanten p_k wenigstens näherungsweise eine periodische Wiederkehr stattfinden, die ihrerseits auf einen unternormalen Wert der Streuung hinwirkt. Da hierbei, wie man sofort übersieht, die Abgrenzung der Versuche nach vollen Jahren wesentlich ist, so kann man fragen, wie sich die Sache gestaltet, wenn die Züge in kürzeren Gruppen, z. B. nach Halb- oder Vierteljahren zusammengefaßt werden. Demgemäß wollen wir im folgenden noch ein Urnenschema betrachten, das man füglich als das allgemeinste bezeichnen kann, weil es imgrunde nicht mehr voraus-

setzt, als daß die Züge in einer gewissen Reihenfolge geordnet vorliegen.

§ 153. Das jetzt zu behandelnde Schema soll folgendermaßen zusammengesetzt werden. Gegeben ist eine Urnenreihe, deren Urnen nicht wie gewöhnlich durch eine einfache Nummer, sondern durch eine Doppelnummer (f, g) unterschieden werden, wobei die Nummer f von 1 bis m, die Nummer g hingegen für jedes f von 1 bis zu einer gewissen von f abhängenden Zahl $n(f)$ laufen soll, so daß, wenn man nach der ersten Nummer gruppenweise zusammenfaßt, die ohne weiteres verständliche Gruppeneinteilung

$$\begin{aligned}
\text{Gruppe } (1) &= (1, 1) + (1, 2) + \cdots + (1, n(1)),\\
\text{Gruppe } (2) &= (2, 1) + (2, 2) + \cdots + (2, n(2)),\\
&\cdots\\
\text{Gruppe } (m) &= (m, 1) + (m, 2) + \cdots + (m, n(m))
\end{aligned}$$

entsteht. Aus jeder Urne wird einmal gezogen, und für jede Gruppe (f) als beobachtetes $x(f)$ die rH. der weißen Züge notiert, die aus den Urnen dieser Gruppe stammen. Gesucht wird die theoretische Verteilung $\mathfrak{W}(x)$ zu der Kollektivreihe, deren m Glieder durch die *Zahlenwerte* der $x(f)$ gebildet werden.

Das angegebene Schema läßt sich, wie man sofort erkennt, aus dem *Poissonschen* Schema ableiten, wenn man voraussetzt, daß die *Poissonsche* Urnenreihe jedesmal, nachdem ein x beobachtet und notiert worden ist, eine gewisse Änderung erleidet, die sowohl die Anzahl der Urnen, als auch ihre Füllungen treffen kann.

Um bei der weiteren Behandlung des Schemas ohne weiteres die in der XIV. Vorlesung entwickelten Sätze über die Mischung von Verteilungen heranziehen zu können, wollen wir zunächst statt der Gruppennummer f den *Index* y durch die Gleichung $y = f : m$ einführen, so daß für y als zulässige Werte die Brüche

$$1 : m, \quad 2 : m, \quad \ldots \quad m : m$$

in Betracht kommen. Entsprechend soll n_y statt $n(f)$ die Urnenmenge in der mit dem Index y bezeichneten Urnengruppe bedeuten. Weiter wollen wir uns, statt von den $x(f)$ auszugehen, an den Umstand halten, daß die Beobachtung m Argumentpaare (x, y) liefert, deren zweiter Bestandteil den Gruppenindex y bedeutet, während der erste Bestandteil x die rH. der weißen Züge anzeigt, die aus der Urnengruppe y herrühren. Diese Auffassung führt uns zu einer aus m Gliedern bestehenden Kollektivreihe, der eine gewisse, von den zwei Argumenten x, y abhängende Verteilung $\mathfrak{U}(x, y)$ zukommt. Dazu kommen dann noch die aus $\mathfrak{U}(x, y)$ abgeleiteten Verteilungen, die wie früher durch die Symbole $\mathfrak{U}(x)$, $\mathfrak{U}(y)$ und $\mathfrak{U}(x)_y$ angezeigt werden sollen. Hierbei ist die Verteilung $\mathfrak{U}(x)$, die die Argumente x ohne Rücksicht auf die

gleichzeitigen Werte von y zusammenfaßt, nichts anderes als das gesuchte $\mathfrak{W}(x)$, da sich ja $\mathfrak{W}(x)$ nur auf die Zahlenwerte der $x(f)$, ohne Berücksichtigung der Gruppennummer, beziehen sollte. Ferner gibt $\mathfrak{U}(y)$ die Verteilung der y, ohne Rücksicht auf die gleichzeitigen Werte von x, an. Da nun innerhalb einer Versuchsreihe die m zulässigen Werte des Index y immer gleich oft, nämlich je einmal, auftreten, so hat der Ausdruck $\mathfrak{U}(y)$ für jedes y denselben Wert, ist also, da die Summe der einzelnen $\mathfrak{U}(y)$ gleich Eins sein muß, gleich $1:m$. Infolgedessen wird der nach $\mathfrak{U}(y)$ genommene Durchschnitt irgend einer von y abhängenden Funktion $T(y)$ gleich dem arithmetischen Mittel aus den $T(y)$, d. h. man hat mit der früher benutzten Schreibweise

$$\mathfrak{D}_y[T(y)] = \sum_y T(y):m. \tag{35}$$

Der Ausdruck $\mathfrak{U}(x)_y$ endlich, der die Mischungsbestandteile für $\mathfrak{U}(x)$ liefert, stellt die Verteilung dar, die dem Argument x bei festgehaltenem y zukommt. Diese Verteilung ergibt sich aber, wenn wir die $\mathfrak{W}$. des weißen Zuges für die Urne (my, g) mit $p(g)_y$ bezeichnen, unmittelbar aus den Formeln des *Poissonschen* Schemas, sobald wir der Rechnung die Größen

$$n_y,\ p(1)_y,\ p(2)_y,\ \ldots\ p(n_y)_y$$

zugrunde legen. Sind solchergestalt die $\mathfrak{U}(x)_y$ festgelegt, so ergibt sich daraus das gesuchte $\mathfrak{U}(x)$ nach § 111 auf Grund der Mischungsgleichung

$$\mathfrak{U}(x) = \mathfrak{D}_y[\mathfrak{U}(x)_y],$$

d. h. $\mathfrak{U}(x)$ ist hier das arithmetische Mittel der m Verteilungen $\mathfrak{U}(x)_y$.

Bei der weiteren Rechnung führen wir noch die reziproken Werte der n_y und ihr Mittel durch die Gleichungen

$$1:n_y = l_y, \qquad l = \mathfrak{D}_y(l_y) \tag{36}$$

ein und behandeln zunächst den Argumentdurchschnitt und die Streuung von $\mathfrak{U}(x)$.

Zu $\mathfrak{U}(x)_y$ gehört der früher mit c_y bezeichnete Argumentdurchschnitt, der jetzt nach dem *Poissonschen* Schema gleich dem arithmetischen Mittel der $p(g)_y$ wird, das mit p_y bezeichnet werden soll. Man erhält also

$$n_y p_y = \sum_g p(g)_y \quad \text{oder} \quad p_y = \sum_g l_y p(g)_y,$$

und damit

$$\mathfrak{D}_x(x) = c_y = p_y = \sum_g l_y p(g)_y. \tag{37}$$

Daraus folgt nach § 112 für den zu $\mathfrak{U}(x)$ gehörigen Durchschnitt, den wir hier mit p bezeichnen wollen,

$$p = \mathfrak{D}(x) = \mathfrak{D}_y(c_y) = \mathfrak{D}_y(p_y) = \mathfrak{D}_y[\sum_g l_y p(g)_y]. \tag{38}$$

Der zu $\mathfrak{U}(x)$ gehörige Argumentdurchschnitt ist also das mit den Gewichten l_y genommene Mittel aller $p(g)_y$, wobei die l_y nach (36) gleich den reziproken Längen der einzelnen Zuggruppen sind.

Die zu einem gegebenen y gehörenden $p(g)_y$ besitzen eine Streuung, deren Quadrat mit P_y bezeichnet werden möge, so daß

$$n_y P_y = \sum_g [p(g)_y - p_y]^2 \quad \text{oder} \quad P_y = \sum_g l_y [p(g)_y - p_y]^2 \tag{39}$$

wird. Setzt man ferner

$$q_y = 1 - p_y, \qquad q = 1 - p,$$

so wird nach der *Poissonschen* Formel die zu $\mathfrak{U}(x)_y$ gehörige Streuung s_y aus

$$n_y s_y^2 = p_y q_y - P_y \quad \text{oder} \quad s_y^2 = l_y(p_y q_y - P_y) \tag{40}$$

gefunden. Ferner gehört zu den p_y eine Streuung, deren Quadrat C heißen möge und durch

$$C = \mathfrak{D}_y [(p_y - p)^2] \tag{41}$$

gegeben ist. Damit erhält man nach § 112 (36) für die zu $\mathfrak{U}(x)$ gehörende Streuung $\operatorname{str}(x)$ die Gleichung

$$\operatorname{str}(x)^2 = C + \mathfrak{D}_y(s_y^2) = C + \mathfrak{D}_y [l_y p_y q_y - l_y P_y]. \tag{42}$$

Zur Umformung dieser Beziehung führen wir noch den Ausdruck

$$L = \mathfrak{D}_y [(l_y - l)(p_y - \tfrac{1}{2})^2] \tag{43}$$

ein, der sich auch in der Gestalt

$$L = \mathfrak{D}_y [(l_y - l)(\tfrac{1}{4} - p_y q_y)]$$

schreiben läßt, woraus wegen (36)

$$L = l \mathfrak{D}_y (p_y q_y) - \mathfrak{D}_y (l_y p_y q_y)$$

folgt. Setzt man nun die Identität

$$p_y q_y = pq + (q - p)(p_y - p) - (p_y - p)^2$$

an und nimmt davon den Durchschnitt nach y, so wird wegen (38) und (41)

$$L = lpq - lC - \mathfrak{D}(l_y p_y q_y). \tag{44}$$

Eliminiert man hiermit die $\mathfrak{D}$-Größe aus (42) und setzt noch

$$\mathfrak{D}_y (l_y P_y) = lP, \tag{45}$$

so entsteht

$$\operatorname{str}(x)^2 = lpq + (1 - l)C - L - lP, \tag{46}$$

worin P offenbar das mit den Gewichten l_y gebildete Mittel der P_y ist.

§ 154. Bei den zu $\mathfrak{U}(x)$ gehörenden Gleichungen (38) und (46) wollen wir stehen bleiben, da die übrigen numerischen Elemente von

$\mathfrak{U}(x)$ auf ziemlich verwickelte Ausdrücke führen, in denen die höheren Potenzmittel der Differenzen

$$p(g)_y - p_y \quad \text{und} \quad p_y - p$$

auftreten. Bei der weiteren Erörterung der gefundenen Formeln dürfen wir uns überdies auf (46) beschränken, da der in (38) gegebene Ausdruck für $\mathfrak{D}(x)$ zu keiner besondern Bemerkung Anlaß gibt. Fallen die Werte der $\mathfrak{W}$-Größen $p(g)_y$ untereinander und folgeweise auch mit den p_y und p zusammen, so verschwinden die Größen C, L, P und man erhält

$$\operatorname{str}(x)^2 = lpq. \tag{47}$$

Wird diese Formel, die sich von dem *Bernoullischen* Schema nur durch die Bedeutung des Faktors l unterscheidet, ähnlich wie früher als Vergleichsnorm genommen, so stellen die von C, L und P abhängenden Terme Abweichungen von der Norm dar. Sind die Größen $p(g)_y$ von Zug zu Zug veränderlich, jedoch so, daß die p_y zusammenfallen, so verschwinden C und L, während die P_y und folgeweise auch P sicher von Null verschieden sind. Die Streuung $\operatorname{str}(x)$ wird dann offenbar unternormal. Sind andrerseits die $p(g)_y$ innerhalb der einzelnen Zuggruppen konstant, während die p_y von Gruppe zu Gruppe schwanken, so verschwinden die P_y nebst P, während C sicher von Null verschieden ist. Die Streuung wird dann sicher übernormal ausfallen, sobald die Gruppenlängen n_y übereinstimmen oder doch nicht stark auseinander gehen, denn in einem solchen Falle wird L verschwinden oder doch wenigstens auf kleine Werte beschränkt bleiben, da ja L von den Schwankungen der l_y abhängt.

Denkt man sich nun weiter noch die verschiedenen Übergänge zwischen den soeben betrachteten Grenzfällen herangezogen, so erkennt man, daß das Verhalten von $\operatorname{str}(x)$, je nach der Gruppierung der Größen $p(g)_y$, eine große Mannigfaltigkeit bieten wird. So kann es im besondern vorkommen, daß die Streuung trotz nicht konstanter $p(g)_y$ den normalen Wert annimmt, indem sich in (46) die von C, L und P abhängenden Terme gegenseitig aufheben. Ein solches Schema wird z. B., wie hier noch gezeigt werden soll, erhalten, wenn man für ein unendliches m und für zusammenfallende n_y die $p(g)_y$ in der Reihenfolge der Züge unter Innehaltung eines gewissen Verteilungsgesetzes $\mathfrak{V}$ nach Zufall variieren läßt. Ist hierbei n der gemeinsame Wert der n_y und demgemäß $l = 1 : n$, so verschwindet zunächst der Ausdruck L. Da ferner unter der gemachten Voraussetzung die n Werte $p(g)_y$ einer Zuggruppe als ebenso viele beobachtete Argumentwerte einer dem Gesetz $\mathfrak{V}$ unterworfenen Kollektivreihe aufgefaßt werden dürfen, so erscheinen die einzelnen p_y als beobachtete Argumentdurchschnitte von $\mathfrak{V}$, und man kann die Streuung dieser Durchschnitte

sofort nach § 126 (19) berechnen. Man erhält danach, wenn s die Streuung von $\mathfrak{V}$ bedeutet, für das Streuungsquadrat C der p_y den Ausdruck

$$C = l s^2. \tag{48}$$

Andrerseits ist jetzt jedes P_y ein an n Gliedern beobachtetes Streuungsquadrat von $\mathfrak{V}$, folglich ist das aus unendlich vielen P_y genommene Mittel P nach § 127 (21) durch den Ausdruck

$$P = (1 - l) s^2 \tag{49}$$

gegeben. Damit nimmt aber (46) die Gestalt

$$\operatorname{str}(x)^2 = l p q \tag{50}$$

an; w. z. b. w.

Der vorstehend betrachtete Fall ist selbstverständlich nicht der einzige, der bei veränderlichen $p(g)_y$ eine normale Streuung zu liefern vermag. Ferner läßt sich bei dem Spielraum, den die Wahl der $p(g)_y$ bietet, leicht ermessen, daß auch bei den Koeffizienten $D_3, D_4, \ldots$ der zu $\mathfrak{U}(x)$ gehörigen Φ-Reihe trotz veränderlicher $p(g)_y$ merklich normale, d. h. mit dem *Bernoullischen* Schema genügend stimmende Werte zustande kommen können. Hiernach wird die oben aufgeworfene Frage nach den Kriterien für die Zulässigkeit des genannten Schemas dahin zu beantworten sein, daß der befriedigende Anschluß einer *einzelnen* beobachteten Verteilungskurve $\mathfrak{U}(x)$ an die *Bernoullische* Formel nicht von vorherein die Veränderlichkeit der $p(g)_y$ ausschließt, weil ja die Übereinstimmung auch durch eine besondere Konfiguration der nicht konstanten p-Größen erzeugt worden sein kann. Anders liegt allerdings die Sache, wenn sich der Anschluß an das normale Schema von einer Versuchsreihe zur andern wiederholt, denn es ist für gewöhnlich nicht sehr wahrscheinlich, daß Konfigurationen der erwähnten Art ständig auftreten. Man wird damit wieder auf die frühere Bemerkung geführt, daß die Menge der Versuche wichtiger ist, als deren Länge.

Achtzehnte Vorlesung.

Schema für seltene Ereignisse und für Gruppen.

§ 155. Läßt man in den Formeln (22) bis (31) der letzten Vorlesung bei festgehaltener Urnenzahl die Werte von p oder q gegen Null gehen, so ergibt sich, daß schließlich die Φ-Reihe unbequem langsam konvergiert, sobald der Wert des Produktes npq unterhalb eines Betrages von wenigen Einheiten liegt. Wenn nun auch hierbei

die Ausdrücke für die numerischen Elemente ihre Bedeutung als charakteristische Bestimmungsstücke der betrachteten Verteilung behalten, so ist es doch wünschenswert, für solche Grenzfälle nicht bloß auf die Benutzung der Φ-Reihe angewiesen zu sein. Deshalb wollen wir die ganze Untersuchung nochmals aufnehmen und dabei zunächst die *Bernoullische* Formel auf demjenigen Wege herleiten, der seither vorzugsweise hierzu benutzt worden ist.

Der Deutlichkeit halber werde zuerst die Aufgabe, um die es sich handelt, wiederholt. Gegeben sind n Urnen mit übereinstimmenden Füllungen; n Züge, je einer aus jeder Urne, bilden zusammen einen Versuch; für jeden Versuch wird als beobachtetes x die rH. der gezogenen weißen Kugeln notiert; gesucht wird die zu x gehörige theoretische Verteilung $\mathfrak{W}(x)$ für den Fall eines *großen* n.

Ist für den einzelnen Zug

$$\mathfrak{W}(\text{weiß}) = p, \qquad \mathfrak{W}(\text{schwarz}) = q, \qquad p + q = 1,$$

so ergeben sich die einzelnen $\mathfrak{W}(x)$ als die Glieder der binomischen Reihe für $(p+q)^n$, und zwar wird mit der Abkürzung $y = 1 - x$

$$\mathfrak{W}(x) = [n! p^{nx} q^{ny}] : [(nx)!\,(ny)!]. \tag{1}$$

Wir untersuchen nun zunächst den Verlauf von $\mathfrak{W}(x)$, soweit er sich unmittelbar aus (1) ablesen läßt, indem wir zugleich die graphische Darstellung von $\mathfrak{W}(x)$ durch die zugehörige Verteilungskurve heranziehen. Hierbei soll vorläufig vorausgesetzt werden, daß weder p noch q klein seien, daß also neben n auch die Produkte np und nq große Werte besitzen.

Setzt man in (1) für nx nacheinander die Zahlen $0, 1, \ldots n$ ein, so erhält man die Reihe der in Betracht kommenden $\mathfrak{W}(x)$. In dieser Reihe dividiere man jedes Glied durch das vorhergehende, so daß die Reihe der Quotienten

$$Q(x) = \mathfrak{W}\left(x + \frac{1}{n}\right) : \mathfrak{W}(x) = pny : [q(nx+1)] \tag{2}$$

entsteht. Die $Q(x)$ nehmen mit wachsendem x beständig ab; ferner ist der erste Quotient gleich $np:q$, also groß und von der Ordnung der Zahl n, andrerseits ist der letzte Quotient gleich $p:nq$, also klein und von der Ordnung der Zahl $1:n$. Danach zerfällt die Reihe der $Q(x)$ in zwei Abschnitte: im ersten Abschnitte sind die Quotienten durchweg größer, im zweiten durchweg kleiner als Eins. Die Grenzscheide liegt an der Stelle, wo $Q(x)$ durch Eins hindurchgeht oder nx den Wert $np - q$ überschreitet. Infolgedessen verläuft die Verteilungskurve anfangs beständig steigend bis zu einem Maximum hin, um nachher beständig zu sinken; das Maximum selber findet in der Umgebung der Stelle $x = p$ statt.

Um über den Abfall der Kurve zu beiden Seiten des Maximums

eine genauere Vorstellung zu gewinnen, betrachten wir zunächst ein numerisches Beispiel. Es sei

$$n = 1000, \quad p = q = 0.5,$$

während für x die Werte von $x' = 0.600$ bis $x'' = 0.610$ gesetzt werden sollen. Man erhält dann

$$Q(0.600) = 400:601, \quad Q(0.609) = 391:610,$$

so daß der Quotient $\mathfrak{W}(x''):\mathfrak{W}(x')$, der gleich dem Produkt der zwischen x' und x'' auftretenden $Q(x)$ ist, sicher kleiner wird als

$$(400:601)^{10} = 1:58.$$

Die Ordinate $\mathfrak{W}(x'')$ ist also kleiner als der 58-ste Teil von $\mathfrak{W}(x')$. Da sich nun diese Betrachtung für andere Werte der Größen n, p, x mit ähnlichem Erfolge wiederholen läßt, so gelangt man zu dem Ergebnis, daß die $\mathfrak{W}(x)$ schon in mäßiger Entfernung vom Maximum auf ganz unmerkliche Beträge herabsinken, daß sich also die Hauptmasse der nach dem Gesetze $\mathfrak{W}(x)$ verteilten x in eine kleine Strecke zu beiden Seiten des Maximums zusammendrängt. Führt man demgemäß statt x und y die neue Veränderliche u durch die Gleichungen

$$2h^2pq = n, \quad u = h(x-p), \tag{3}$$

$$x = p + \frac{u}{h}, \quad y = q - \frac{u}{h} \tag{4}$$

ein, so darf man sich nach dem Gesagten, so lange nur jene Hauptmasse der x untersucht werden soll, auf mäßige Werte von u beschränken, da ja das Maximum von $\mathfrak{W}(x)$ bei der Stelle $x = p$ oder $u = 0$ liegt.

§ 156. Nach den vorstehenden Bemerkungen formen wir jetzt den Ausdruck (1) um, indem wir die in § 29 (29) gegebene Darstellung der Funktion $\Pi(X)$, nämlich die Gleichung

$$\log \Pi(X) = X \log X - X + \tfrac{1}{2}\log(2\pi X) + J \tag{5}$$

heranziehen, in der das Glied J mit wachsendem X gegen Null geht. Bildet man nach (1) den Logarithmus von $\mathfrak{W}(x)$ und setzt darin nach (5) die Ausdrücke für die Logarithmen von $n!$, $(nx)!$ und $(ny)!$ ein, so ergibt sich eine Gleichung von der Gestalt

$$\log \mathfrak{W}(x) = nx(\log p - \log x) + ny(\log q - \log y)$$
$$- \tfrac{1}{2}\log(xy) + K + N,$$

wo unter K alle nach x konstanten Glieder zusammengefaßt sind, während N nur solche Bestandteile enthält, welche bei festgehaltenen Werten von p, q, x, y für ein unendliches n verschwinden. Führt man dann statt x und y nach (3) und (4) die Veränderliche u ein und entwickelt nach Potenzen von u, so wird

$$\log \mathfrak{W}(x) = -u^2 + K' + N',$$

wo K' und N' eine ähnliche Bedeutung haben, wie vorhin K und N. Unterdrückt man N', so erhält man $\mathfrak{W}(x)$ in der Gestalt

$$\mathfrak{W}(x) = A \exp(-u^2),$$

wo A eine gewisse Konstante bedeutet. Damit ist man in Wahrheit bereits bei dem einfachen E.-G. angelangt, und es würde, um die *Bernoullische* Formel zu erhalten, nur noch nötig sein, den Übergang von der Verteilung $\mathfrak{W}(x)$ auf die Summenfunktion und entsprechend von $\exp(-u^2)$ auf die Funktion Φ auszuführen. Wir wollen uns jedoch nicht dabei aufhalten, weil das Vorstehende nur ausgeführt worden ist, um der Vollständigkeit halber eine Vergleichung des herkömmlichen Verfahrens mit den Methoden der letzten Vorlesung zu geben. Der hier skizzierte Weg läßt sich, wie noch bemerkt werden mag, durch gewisse Kunstgriffe in eine ziemlich elementare Gestalt bringen und ist dann da sehr gut am Platze, wo man absichtlich auf die Benutzung weiter gehender analytischer Hilfsmittel verzichten will. Mit Rücksicht darauf ist denn die *Bernoullische* Formel gelegentlich auch geradezu als „Binomialsatz" bezeichnet worden. Im übrigen erkennt man, daß es einigermaßen schwierig sein würde, auf dem betrachteten Wege das Bildungsgesetz der in der *Bernoullischen* Formel vernachlässigten Glieder in übersichtlicher Gestalt darzustellen.

Da vorläufig vorausgesetzt worden war, daß np dieselbe Größenordnung wie n besitze, so ist jetzt zu untersuchen, wie sich $\mathfrak{W}(x)$ umgestaltet, wenn np nicht mehr groß ist. Läßt man nun bei festgehaltenem n die Größe p abnehmen, so rückt in der vorhin betrachteten Verteilungskurve das Maximum nach der Stelle $x = 0$ hin und fällt schließlich in die erste Ordinate, sobald der erste Quotient $Q(x)$ oder der Bruch $np:q$ kleiner als Eins geworden ist. Hierbei bleibt aber die Bemerkung über den Abfall der Kurve $\mathfrak{W}(x)$ bestehen, d. h. die Hauptmasse der x drängt sich wiederum in der Umgebung des Maximums zusammen. Bildet man daher

$$np = a, \qquad q = 1 - \frac{a}{n}, \qquad nx = z, \qquad n - nx = n - z$$

und setzt für a mäßige Beträge voraus, so sind nach dem Gesagten bei der auf ganze Zahlen beschränkten Größe z ebenfalls nur mäßige Werte zu berücksichtigen. Demgemäß setzen wir nach (1) und (5) unter Vernachlässigung des Bestandteiles J in (5) die Gleichung

$$\begin{aligned}\log \mathfrak{W}(x) = n \log n - n + \tfrac{1}{2}\log(2\pi n)\\ -(n-z)\log(n-z) + (n-z) - \tfrac{1}{2}\log(2\pi n - 2\pi z)\\ + z\log\frac{a}{n} + (n-z)\log\left(1-\frac{a}{n}\right) - \log \Pi(z)\end{aligned}$$

an und entwickeln darauf die Ausdrücke

$$\log(n-z) \quad \text{und} \quad \log\left(1-\frac{a}{n}\right)$$

nach Potenzen von z und a. Daraus ergibt sich, wenn die für ein unendliches n verschwindenden Glieder unterdrückt werden, nach der nötigen Reduktion

$$\log \mathfrak{W}(x) = z \log a - a - \log \Pi(z).$$

Bezeichnet man noch, da kein Mißverständnis zu befürchten ist, die Verteilung von z kurz durch $\mathfrak{W}(z)$, so wird

$$\mathfrak{W}(z) = a^z : [\Pi(z) \exp a], \tag{6}$$

wo z die Menge der bei dem Versuche gezogenen weißen Kugeln bedeutet. Die vorstehende Gleichung enthält das *Urnenschema für seltene Ereignisse* und ist im Falle eines kleinen p an die Stelle der *Bernoullischen* Näherungsformel zu setzen. Es läßt sich nun zeigen, daß die gefundene Gleichung auch noch bestehen bleibt, wenn man den *Poissonschen* Urnenansatz statt des *Bernoullischen* zugrunde legt.

§ 157. Wenn die gegebenen Urnen ungleiche Füllungen enthalten, so sei wieder wie früher für den einzelnen Zug aus der k-ten Urne

$$\mathfrak{W}(\text{weiß}) = p_k, \qquad \mathfrak{W}(\text{schwarz}) = q_k, \qquad p_k + q_k = 1.$$

Ferner bedeute z die Menge der bei einem Versuche gezogenen weißen Kugeln und $\mathfrak{W}(z)$ die theoretische Verteilung der z. Dann besteht, wie wir aus § 147 (10) wissen, die nach $\mathfrak{W}(z)$ gebildete Durchschnittsgleichung

$$\mathfrak{D}(u^z) = (up_1 + q_1)(up_2 + q_2) \cdots (up_n + q_n). \tag{7}$$

Die p_k sollen jetzt kleine Zahlen von der Größenordnung $1 : n$ sein. Demgemäß setzen wir

$$np_k = a_k, \qquad n a = \sum_k a_k,$$

wo der Voraussetzung nach die a_k und ihr arithmetisches Mittel a auf mäßige Beträge beschränkt bleiben. Entwickelt man nun die Logarithmen

$$\log (up_k + q_k) = \log \left[1 + \frac{u-1}{n} a_k\right]$$

nach Potenzen von $u - 1$ und summiert darauf die entstehenden Reihen nach k, so nimmt das erste Glied der Summe die Gestalt $a(u-1)$ an. Das zweite Glied, das wir ebenfalls berücksichtigen wollen, schreiben wir in der Gestalt $-NA(u-1)^2$, wo

$$N = 1 : n, \qquad 2An = \sum_k a_k^2$$

ist, und A die Größenordnung von a^2 besitzt. Die übrigen Glieder vernachlässigen wir und erhalten

$$\log \mathfrak{D}(u^z) = a(u-1) - NA(u-1)^2,$$

$$\mathfrak{D}(u^z) = \exp(au - a) \cdot \exp(-NAu^2 + 2NAu - NA). \tag{8}$$

Setzt man die Reihenentwicklung

$$\exp(au - a) = w(0) + w(1)u + w(2)u^2 + \cdots \tag{8.a}$$

an, so wird für $z = 0, 1, 2, \ldots$

$$w(z) = a^z : [z! \exp a]. \tag{9}$$

Beachtet man nun, daß

$$\mathfrak{D}(u^z) = \mathfrak{W}(0) + \mathfrak{W}(1)u + \mathfrak{W}(2)u^2 + \cdots \tag{9.a}$$

ist, und unterdrückt zunächst in (8) die von N abhängenden Glieder, so entsteht die für hinreichend große n genügende Näherung

$$\mathfrak{W}(z) = w(z), \tag{10}$$

die mit (6) übereinstimmt, nur daß jetzt a das arithmetische Mittel aus den für die einzelnen Urnen geltenden a_k bedeutet. Berücksichtigt man in (8) dagegen noch die Glieder, die nach N von der ersten Ordnung sind, so erhält man aus (8)

$$\mathfrak{D}(u^z) = \exp(au - a)[1 - NAu^2 + 2NAu - NA],$$

oder, wenn man für den Exponentialfaktor die Reihe (8.a) einsetzt,

$$\mathfrak{D}(u^z) = [w(0) + w(1)u + \cdots] \cdot [1 - NAu^2 + 2NAu - NA].$$

Entwickelt man nun rechts nach u und vergleicht mit der in (9.a) angesetzten Reihe für $\mathfrak{D}(u^z)$, so ergibt sich

$$\mathfrak{W}(z) = w(z) - NA[w(z) - 2w(z-1) + w(z-2)]. \tag{11}$$

Der Gebrauch dieser Formel wird offenbar wesentlich erleichtert, wenn man über Tafeln verfügt, die den Wert von $w(z)$ mit den Argumenten z und a unmittelbar zu entnehmen gestatten.

§ 158. Kehrt man wieder zu der abgekürzten Formel (10) oder dem Ansatze $N = 0$ zurück, so folgt aus (8)

$$\mathfrak{D}(u^z) = \exp(au - a). \tag{12}$$

Differentiiert man diese Gleichung wiederholt nach u und setzt darauf $u = 1$, so wird

$$\mathfrak{D}[z] = a,$$
$$\mathfrak{D}[z(z-1)] = a^2,$$
$$\mathfrak{D}[z(z-1)(z-2) = a^3,$$
usw.

Die Summe der beiden ersten Gleichungen liefert

$$\mathfrak{D}(z^2) = a + a^2.$$

Da nun allgemein $\operatorname{str}(z)^2 = \mathfrak{D}(z^2) - \mathfrak{D}(z)^2$ ist, so erhält man für die beiden ersten numerischen Elemente von $w(z)$ die Ausdrücke

$$\mathfrak{D}(z) = a, \qquad \operatorname{str}(z)^2 = a. \tag{13}$$

Die Gleichung (10), die sich bei *Poisson* (*Recherches sur la probabilité des jugements* S. 206) findet, enthält die für seltene Ereignisse zweckmäßige Umgestaltung des in § 149 (32) gegebenen *Poissonschen* Urnenschemas. Es ist bemerkenswert, daß $w(z)$ nur einen einzigen Parameter enthält, indem die übrigen in $\mathfrak{W}(z)$ tatsächlich vorkommenden Parameter nur in solchen Gliedern auftreten, die für ein hinreichend großes n unmerklich werden. Tafeln für $w(z)$ finden sich in einer Untersuchung von *L. von Bortkewitsch* „*Das Gesetz der kleinen Zahlen*" (Leipzig 1898). Ebenda sind auch Vergleichungen mit verschiedenen hierher gehörenden Beobachtungsreihen, z. B. Tod durch Unfall oder Selbstmord, ausgeführt; die dabei auftretenden Widersprüche zwischen Beobachtung und Rechnung verlaufen im allgemeinen befriedigend und sind in einzelnen Fällen sogar überraschend klein.

Geht man noch zu der in § 153 behandelten Erweiterung des *Poissonschen* Urnenschemas über und bezeichnet mit $\mathfrak{W}(z, f)$ die Verteilungsfunktion, die zur f-ten Urnengruppe gehört, so ist die Funktion $\mathfrak{W}(z)$, die aus der Zusammenfassung der m betrachteten Urnengruppen entspringt, gleich dem arithmetischen Mittel der $\mathfrak{W}(z, f)$, also

$$m\mathfrak{W}(z) = \mathfrak{W}(z, 1) + \mathfrak{W}(z, 2) + \cdots + \mathfrak{W}(z, m). \tag{14}$$

Bedeutet a_f den Argumentdurchschnitt zu $\mathfrak{W}(z, f)$, so erhält man für die beiden ersten Elemente von $\mathfrak{W}(z, f)$ nach (13) die Ausdrücke

$$\mathfrak{D}(z) = a_f, \qquad \operatorname{str}(z)^2 = a_f.$$

Ist ferner a das arithmetische Mittel der a_f, und C ihr Streuungsquadrat, also

$$mC = \sum_f (a_f - a)^2,$$

so ergeben sich für die beiden ersten Elemente der gemischten Verteilung $\mathfrak{W}(z)$ nach der Mischungsregel die Ausdrücke

$$\mathfrak{D}(z) = a, \qquad \operatorname{str}(z)^2 = a + C. \tag{15}$$

Die Streuung wird also übernormal, sobald die a_f ungleich sind. Dieses Ergebnis hätte man übrigens aus § 153 (46) auch direkt durch einen Grenzübergang herleiten können.

§ 159. Wenn a sehr klein ist, so wird, wie bereits bemerkt wurde, die Φ-Reihe für die numerische Rechnung ungeeignet. Sie wird jedoch wieder benutzbar, sobald der Wert von a einige Einheiten übersteigt; deshalb mögen hier noch die ersten D-Koeffizienten der Verteilung $w(z)$ abgeleitet werden.

Formt man die Beziehung

$$\mathfrak{D}(u^z) = \exp(au - u)$$

vermittelst der Gleichungen

$$2ab^2 = 1, \qquad u = \exp(-2bv) \tag{16}$$

um, so entsteht zunächst

$$\mathfrak{D}[\exp(-2bzv)] = \exp[a\exp(-2bv) - a].$$

Entwickelt man auf der rechten Seite innerhalb [ ] nach v und setzt für den Augenblick

$$A = \exp[-2abv + v^2],$$

$$B = \exp\left[-\frac{b(2v)^3}{2\cdot 3!} + \frac{b^2(2v)^4}{2\cdot 4!} - \frac{b^3(2v)^5}{2\cdot 5!} + \cdots\right],$$

so wird unter Beachtung der Beziehung zwischen a und b

$$\mathfrak{D}[\exp(-2bzv)] = A\cdot B.$$

Zieht man jetzt den Faktor A auf die andere Seite und führt das Hilfsargument

$$u = b(z-a)$$

ein, so entsteht

$$B = \mathfrak{D}[\exp(-2uv - v^2)] = \sum_p \mathfrak{D}[\mathfrak{R}(u)_p](2v)^p.$$

Man erhält also sofort die gesuchten D-Koeffizienten, wenn man B nach Potenzen von $2v$ entwickelt. Damit wird

$$D_0 = 1, \qquad D_1 = 0, \qquad D_2 = 0,$$

$$12D_3 = -b, \qquad 48D_4 = b^2, \tag{17.a}$$

$$240D_5 = -b^3, \qquad 1440D_6 = b^4 + 5b^2. \tag{17.b}$$

Da D_1 und D_2 verschwinden, so sind a und b die normalen Werte der sonst mit c und h bezeichneten Parameter.

§ 160. Anschließend an die Betrachtung der Verteilung seltener Ereignisse soll jetzt noch die Frage der sogenannten Sequenzen behandelt werden. Man denke sich, daß aus einer Urne, die weiße und schwarze Kugeln enthält, fortlaufend unter jedesmaliger Zurücklegung der Kugel gezogen werde, und daß für den einzelnen Zug

$$\mathfrak{W}(\text{weiß}) = p, \qquad \mathfrak{W}(\text{schwarz}) = q$$

sei. Denkt man sich ferner zunächst die Reihe der Zugnummern 1, 2, ... hingeschrieben und darauf jede Zugnummer durch den Buchstaben p oder q ersetzt, je nachdem der betreffende Zug eine weiße oder schwarze Kugel zum Vorschein brachte, so wird auf diese Weise die beobachtete Zugreihe samt ihren Farbenfolgen und Farbenwechseln übersichtlich durch eine aus den Zeichen p und q zusammengesetzte Buchstabenreihe dargestellt. Greift man aus dieser Reihe r aufeinander folgende Buchstaben heraus, so erhält man eine r-gliedrige Zuggruppe, die kurz mit $G(r)$ bezeichnet werden möge. Hierbei kann es dann vorkommen, daß die betrachtete Gruppe Kugeln von nur einer Farbe enthält. Tritt dieser Fall ein, so wollen wir die Gruppe als

eine *Sequenz* $S(r)$ bezeichen. Die W. für das Eintreten einer solchen Sequenz ist offenbar p^r oder q^r, je nachdem es sich dabei um weiße oder schwarze Kugeln handelt.

Nimmt in der Sequenz $S(r)$ die Gliederzahl beständig zu, so nimmt die zugehörige W. beständig ab. Dasselbe gilt aber auch für jede Gruppe $G(r)$, sobald man darin für die Buchstaben p, q eine *bestimmte* Reihenfolge vorschreibt. Wenn z. B. $p = q = 0.5$ ist, wie das u. a. für das Spiel „Bild oder Schrift" bei normaler Beschaffenheit der geworfenen Münze zutrifft, so ist für jede Gruppe $G(r)$ mit vorgeschriebener Anordnung der Buchstaben der Wert der W. ebenso groß wie für die Sequenzen $S(r)$, nämlich gleich $1:2^r$. Es besteht also nach den Ansätzen der W.-R. zwischen den $G(r)$ und $S(r)$ kein wesentlicher Unterschied. Diese Gleichstellung ist nun aber nicht unangefochten gebliebeu. In den Schriften von *d'Alembert* begegnet man an verschiedenen Stellen dem Versuche, die These zu erhärten, daß die $S(r)$ oberhalb eines gewissen r zwar logisch möglich, dagegen physisch unmöglich seien, und daß man solche Sequenzen, wie man sie bisher nicht beobachtet habe, auch in Zukunft nicht beobachten werde; [vgl. beispielsweise die Ausführungen von *d'Alembert* in Tome II (Seite 10) seiner „*Opuscules mathématiques*" (Paris 1761)]. Da er die physische Möglichkeit von Gruppen $G(r)$, welche die Buchstaben p, q in vorgeschriebener Anordnung gemischt enthalten, nicht leugnet, sondern einfach beiseite läßt, so läuft seine These offenbar darauf hinaus, daß den Sequenzen eine grundsätzlich andere Stellung zuzuweisen sei, als den „gemischten" Gruppen. Diese Anschauung ist nun in der Folgezeit bis vor kurzem immer abgelehnt worden, und konnte auch garnicht zugelassen werden, so lange man nicht die Grundvoraussetzung aller Entwicklungen der W.-R., nämlich die Unabhängigkeit der einzelnen Züge des Urnenschemas, aufgeben wollte. Man muß also, wenn man einmal die Ansicht *d'Alemberts* ernsthaft diskutieren will, jene Voraussetzung fallen lassen und darauf zurückgreifen, daß die stets angenommene Unabhängigkeit im *strengsten* Sinne des Wortes tatsächlich nicht vorhanden ist, und daß infolgedessen zwischen den Urnenzügen oder — um einen weniger schematischen Fall zu nehmen — beispielsweise zwischen den einzelnen Würfen bei dem Spiel „Bild oder Schrift" eine gewisse Verbundenheit bestehen könne, die gegen das Auftreten der höheren Sequenzen wirkt. Da man nun aber den Verlauf der Verbundenheit, wenn man sie einmal zuläßt, nicht kennt und deshalb ihre Wirkungsweise nicht mathematisch verfolgen kann, so bleibt nur der Weg des Experiments übrig, um zu prüfen, ob die tatsächliche Häufigkeit der höheren Sequenzen hinter den Beträgen zurückbleibt, die aus dem üblichen Ansatz der W.-R. entspringen. Diesen Weg hat *K. Marbe* in einer Abhandlung „*Naturphilosophische Untersuchungen zur Wahrscheinlichkeitslehre*" (Leipzig 1899) betreten,

mit der ausgesprochenen Absicht, die Meinung von *d'Alembert* wieder zu Ehren zu bringen. *Marbe* zieht hierzu die Aufzeichnungen über rund 80000 ausgeführte Roulettespiele heran und glaubt durch seine Rechnungen den von ihm beabsichtigten Nachweis erbracht zu haben. Es ist hier nicht nötig, auf die *Marbeschen* Deduktionen, die von Anfang an sehr kühl aufgenommen wurden und in mathematischer Hinsicht teilweise verfehlt sind, im einzelnen einzugehen, es genügt vielmehr für unsern Zweck, von dem Ergebnis der Auszählungen als wesentlich das viermalige Vorkommen von $S(12)$ und das Fehlen der höheren Sequenzen festzuhalten und zu fragen, ob dieses Verhalten des Spiels einen deutlichen Widerspruch gegen den üblichen Ansatz der W.-R. in sich schließe.

§ 161. Da die Roulettescheibe außer einem Nullfelde je 18 rote und schwarze Felder enthält, so hat man, um das Urnenschema den Spielbedingungen anzupassen, die Urnenfüllung in der Weise abzuändern, daß sie außer einer mit Null bezeichneten Kugel noch je 18 rote und schwarze Kugeln enthält. Bezeichnet man nun den Bruch $18:37$ mit P, so ist die W. der roten Sequenz $S(r)$ gleich P^r, und das Gleiche gilt von der schwarzen Sequenz $S(r)$. Folglich ist $2P^r$ die W., daß überhaupt eine Sequenz $S(r)$, sei es rot oder schwarz, eintrete. Denkt man sich weiter r Züge ausgeführt und mit x die Anzahl der dabei beobachteten Sequenzen bezeichnet, so ist x entweder Null oder Eins. Wird der angegebene Versuch m-mal ausgeführt, so entsteht eine Kollektivreihe von dem Umfange m, bei der als mögliche Argumente die Werte $x = 0$ und $x = 1$ mit einer gewissen Verteilung $U(x)$ auftreten, deren theoretischer Verlauf durch

$$\mathfrak{U}(1) = 2P^r, \quad \mathfrak{U}(0) = 1 - 2P^r$$

gegeben ist. Multipliziert man die U-Größen mit dem Umfange m, so erhält man statt der relativen Häufigkeiten die absoluten, so daß der Ausdruck

$$m\mathfrak{U}(1) = 2mP^r \tag{18}$$

die Zahl liefert, die unmittelbar mit der beobachteten Menge der $S(r)$ zu vergleichen ist.

Die Menge der in der Kollektivreihe vereinigten Züge ist mr. Legt man für diese Zahl mit Rücksicht auf die von *Marbe* benutzten Aufzeichnungen den runden Betrag 80000 zugrunde und rechnet für einige Werte von r die entsprechenden Werte von $m\mathfrak{U}(1)$ nach (18), so erhält man in dem nachstehenden Täfelchen die drei ersten Zeilen, die unmittelbar verständlich sind.

$r =$	12	13	14	15
$mU(1) =$	4	0	0	0
$m\mathfrak{U}(1) =$	2.34	1.05	0.48	0.22
$\text{str}(U) =$	1.53	1.02	0.69	0.46

Da es nun bei der Beurteilung der Widersprüche zwischen Beobachtung und Rechnung, d. h. also hier der Differenzen $mU - m\mathfrak{U}$, auch auf die der Beobachtung anhaftenden Streuungen ankommt, so sind nach § 130 (31) die in der letzten Zeile des Täfelchens angesetzten Streuungen vermittelst der Gleichung

$$\mathrm{str}\,(U)^2 = m\mathfrak{U}(1 - \mathfrak{U})$$

berechnet worden.

Ein Blick in das Täfelchen lehrt, daß die Widersprüche zwischen Beobachtung und Rechnung von derselben Größenordnung sind, wie die Streuungen, daß also das Ausbleiben der Sequenzen von $r = 13$ ab in der vorgelegten Reihe schlechterdings nichts auffälliges bietet und darum auch in keiner Weise als Stütze für die Anschauung *d'Alemberts* herangezogen werden kann. Anders läge ja die Sache, wenn m 100-mal größer genommen worden wäre, denn dann würden die drei letzten Spalten, falls die beobachteten Sequenzen wieder nur bis $S(12)$ gingen, folgendes Aussehen zeigen:

$r =$	13	14	15
$mU(1) =$	0	0	0
$m\mathfrak{U}(1) =$	105	48	22
$\mathrm{str}\,(U) =$	10.2	6.9	4.6

Man hätte in diesem Falle offenbar guten Grund, von einem auffälligen Widerspruch zwischen Beobachtung und Rechnung zu sprechen. Ebenso hätte man aber auch Grund zu sagen, daß die Zeit, die das Auszählen von 8 Millionen Roulettespielen kosten würde, für andere Dinge nützlicher verwendet werden könnte. Denn der Streit um die These von *d'Alembert* ist imgrunde genommen müßig, weil sie durch bloße logische Deduktionen aus feststehenden Vordersätzen weder allgemein bewiesen, noch allgemein widerlegt werden kann, und weil ferner den Verteidigern der These, wenn sie in einem konkreten Falle, wie ihn das hier benutzte Beispiel bietet, zurückgewiesen werden, immer noch der Rückzug in das Gebiet der höheren Sequenzen offen bleibt, wohin das Experiment nicht nachfolgen kann.

§ 162. Die vorhin angestellte Rechnung beruhte darauf, daß man durch Zusammenfassen der Zugnummern 1 bis r, $r + 1$ bis $2r$, usw. die Gruppen $G(r)$ bildet, die der Abzählung der Sequenzen $S(r)$ zugrunde zu legen sind. Ein solches Abteilen der ganzen Zugreihe in *getrennte* Gruppen ist notwendig, wenn man die Formeln des *Bernoulli*schen Schemas benutzen will, da ja dieses Schema aus Versuchen zusammengesetzt ist, die keinen Zug gemeinsam haben. Hierbei werden nun aber alle diejenigen Sequenzen $S(r)$ zerschnitten und folgeweise auch nicht mitgezählt, deren letzte Zugnummer nicht durch r teilbar ist; ferner bleiben, wenn die Zügezahl n ebenfalls

nicht durch r teilbar ist, gewisse Züge, die dem Ende der Reihe angehören, überhaupt unbenutzt. Will man dem entgehen, so hat man das ganze Schema umzugestalten, und zwar derart, daß *alle* Gruppen $G(r)$, die sich aus der vorgelegten Zugreihe bilden lassen, gleichmäßig herangezogen werden. Hierbei liegt es dann nahe, die Aufgabe zu erweitern, d. h. nicht bloß die Sequenzen zu berücksichtigen, sondern ganz allgemein nach der Verteilung zu fragen, die einer oder auch mehreren bestimmten Gruppenformen zukommt. Als Beispiel hierzu möge der im folgenden behandelte Fall der zweigliedrigen Gruppen dienen.

Als Ausgangspunkt diene wieder die Urne mit weißen und schwarzen Kugeln und den zugehörigen $\mathfrak{W}$-Größen

$$\mathfrak{W}(\text{weiß}) = p, \quad \mathfrak{W}(\text{schwarz}) = q. \tag{19}$$

Aus der Urne wird fortlaufend unter Zurücklegung der Kugel gezogen und die Reihe der Gruppen $G(2)$ gebildet, die dadurch entstehen, daß man die Zugnummern nach dem Schema

$$(1, 2), \quad (2, 3), \quad (3, 4), \text{ usw.}$$

zusammenfaßt. Benutzt man ferner wie oben die Buchstaben p und q zur Bezeichnung der weißen und der schwarzen Züge, so gehört jedes Zugpaar einer der vier Formen

$$(pp), \quad (pq), \quad (qp), \quad (qq)$$

an, und man hat, wenn die Auszählung auf die n ersten Züge mit den dazu gehörigen $n - 1$ Gruppen $G(2)$ beschränkt bleibt, zunächst festzusetzen, welche Formen gezählt oder nicht gezählt werden sollen. So sind z. B. für die Untersuchung der Sequenzen $S(2)$ nur die Formen (pp) und (qq), sei es getrennt, sei es vereint, zu berücksichtigen. Dem entsprechend mögen die Zugpaare, die bei der Auszählung jedesmal berücksichtigt werden sollen, kurz als die *zählenden* Gruppen oder Formen bezeichnet werden.

Bedeuten x und y die Mengen der zählenden und der nichtzählenden Gruppen in der beobachteten Zugreihe, so ist immer $x + y = n - 1$. Da hiernach die beiden Argumente x und y durch eine einfache lineare Transformation zusammenhängen, so läßt sich die Verteilung von y sofort hinschreiben, wenn man die Verteilung von x kennt. Folglich kann man, da die Zählung von vier Formen von vornherein außer Betracht bleibt, zunächst die Zählung von drei Formen beiseite lassen, da sie durch die Fälle mit nur einer zählenden Form bereits erledigt ist. Ferner darf man von den sechs Fällen mit zwei zählenden Formen, nämlich

$$(pp) \text{ und } (pq), \quad (pp) \text{ und } (qp), \quad (pp) \text{ und } (qq),$$
$$(pq) \text{ und } (qp), \quad (pq) \text{ und } (qq), \quad (qp) \text{ und } (qq),$$

die drei letzten beiseite lassen, da sie durch die drei ersten erledigt sind. Endlich gehen die Fälle mit nur einer zählenden Form ineinander über, wenn man die Größen p und q ihre Rolle tauschen läßt. Demnach darf man die Untersuchung auf die nachstehenden fünf Fälle beschränken:

I.	Zählende	Formen:	nur (pp),
II.	„	„	: nur (pq),
III.	„	„	: (pp) und (qq),
IV.	„	„	: (pp) und (pq),
V.	„	„	: (pp) und (qp).

Die verschiedenen Fälle lassen sich, wie wir sehen werden, vorläufig gemeinsam behandeln und sind erst später zu trennen.

§ 163. Der k-te Zug aus der Urne werde, solange es auf das besondere Ergebnis des Zuges nicht ankommt, mit x_k bezeichnet. Sollen dagegen die beiden möglichen Zugergebnisse „weiß“ und „schwarz“ auseinander gehalten werden, so schreiben wir p_k oder q_k statt x_k. Zu der Veränderlichen x_k gehört eine Verteilungsfunktion $\mathfrak{U}(x_k)$, deren Verlauf durch die beiden Gleichungen

$$\mathfrak{U}(p_k) = p, \quad \mathfrak{U}(q_k) = q \tag{20}$$

bestimmt ist.

Weiter werde das Zeichen $e(x_k x_{k+1})$ eingeführt, das die Zahl Null oder Eins bedeuten soll, je nachdem das Zugpaar $(x_k x_{k+1})$ eine zählende oder nichtzählende Gruppe liefert. Demnach ist, wenn k oberhalb $n-1$ liegt, beständig $e = 0$, da im vorliegenden Falle überhaupt nur die Gruppen innerhalb der n ersten Züge zählen sollen. Dagegen richtet sich für ein unterhalb n gelegenes k der Wert der e-Größen jedesmal danach, welcher der fünf oben genannten Fälle gerade vorliegt. Dies festgesetzt gibt der Ausdruck

$$x = e(x_1 x_2) + e(x_2 x_3) + \cdots \tag{21}$$

die Menge der zählenden Gruppen innerhalb der aus n Zügen bestehenden Zugreihe an, und die Aufgabe besteht jetzt darin, die zu dem Argument x gehörende Verteilungsfunktion aufzusuchen, die unter Berücksichtigung der Zügezahl n mit $\mathfrak{U}(x, n)$ bezeichnet werden soll. Zur Lösung der Aufgabe soll der Ausdruck

$$S(n) = \mathfrak{D}(u^x, n) = \sum_x \mathfrak{U}(x, n) u^x \tag{22}$$

dienen, wo das Zeichen $\mathfrak{D}(u^x, n)$ den nach $\mathfrak{U}(x, n)$ genommenen Durchschnitt von u^x bedeutet, und wo ferner die Summation nach x von 0 bis ∞ laufen darf, da ja die $\mathfrak{U}(x, n)$ für die außermöglichen x von selber verschwinden.

Da x durch (21) als Funktion der voneinander unabhängigen

Veränderlichen x_k definiert ist, so liegt eine Argumentmischung vor, und man erhält mit der Substitution

$$u = \exp v$$

aus (22) die Beziehung

$$S(n) = \mathfrak{D}[\exp[ve(x_1x_2) + ve(x_2x_3) + \cdots]], \tag{23}$$

in der auf der rechten Seite die $\mathfrak{D}$-Operation nach den in (20) aufgestellten Verteilungen der sämtlichen x_k auszuführen ist. Hierbei ist zu beachten, daß wegen (20) für eine beliebige von x_k abhängende Funktion $F(x_k)$ der nach x_k genommene Durchschnitt die Gestalt

$$\begin{aligned}\mathfrak{D}(F) &= \mathfrak{U}(p_k)F(p_k) + \mathfrak{U}(q_k)F(q_k) \\ &= pF(p_k) + qF(q_k)\end{aligned} \tag{24}$$

besitzt. Um nun diese Formel bei der schrittweisen Ausführung der in (23) geforderten $\mathfrak{D}$-Operationen bequemer anwenden zu können, sollen zunächst einige Abkürzungen eingeführt werden. Man denke sich das in (23) auftretende Aggregat

$$ve(x_1x_2) + ve(x_2x_3) + \cdots$$

von den Bestandteilen befreit, die aus den $r - 1$ ersten Zügen entspringen und ersetze darauf die nunmehr an erster Stelle stehende Veränderliche x_r durch p_r oder q_r. Dann erhält man statt der Größe $S(n)$ gewisse andere Ausdrücke, die wir in der Gestalt

$$[p, r, n] = \mathfrak{D}[\exp[ve(p_rx_{r+1}) + ve(x_{r+1}x_{r+2}) + \cdots]], \tag{25.a}$$

$$[q, r, n] = \mathfrak{D}[\exp[ve(q_rx_{r+1}) + ve(x_{r+1}x_{r+2}) + \cdots]] \tag{25.b}$$

ansetzen wollen, worin das $\mathfrak{D}$-Zeichen wie bisher auf alle dahinter stehenden Veränderlichen x_k zu beziehen ist. Schreibt man ferner in (25.a) p_{r+1} statt x_{r+1}, so entsteht eine neue $\mathfrak{D}$-Größe, aus der sich der konstante Faktor

$$\exp[ve(p_rp_{r+1})]$$

herausziehen läßt. Ähnliches tritt ein, wenn man von (25.b) ausgeht oder aber x_{r+1} durch q_{r+1} ersetzt. Wir wollen diese heraustretenden konstanten Faktoren durch die Abkürzungen

$$\begin{aligned}\exp[ve(p_rp_{r+1})] &= y(pp)_r, \quad \exp[ve(p_rq_{r+1})] = y(pq)_r, \\ \exp[ve(q_rp_{r+1})] &= y(qp_r), \quad \exp[ve(q_rq_{r+1})] = y(qq)_r,\end{aligned} \tag{26}$$

bezeichnen. Nach diesen Festsetzungen werde jetzt auf der rechten Seite von (25.a) die $\mathfrak{D}$-Operation nach x_{r+1} unter Berücksichtigung von (24) ausgeführt. Dann entsteht wegen (26) die Relation

$$[p, r, n] = py(pp)_r[p, r + 1, n] + qy(pq)_r[q, r + 1, n]. \tag{27.a}$$

In derselben Weise folgt aus (25.b)

$$[q, r, n] = py(qp)_r[p, r + 1, n] + qy(qq)_r[q, r + 1, n]. \tag{27.b}$$

Andrerseits liefert derselbe Prozeß, an (23) nach x_1 ausgeführt, die Gleichung

$$S(n) = \mathfrak{D}(u^x, n) = p[p, 1, n] + q[q, 1, n]. \tag{28}$$

Da die Größen $e(x_r x_{r+1})$ beständig null sind, wenn die Nummer r den Wert n erreicht oder überschreitet, so erhält man aus (25) für $r = n$

$$[p, n, n] = 1, \quad [q, n, n] = 1. \tag{29}$$

Man kann also nach (27), von $r = n - 1$ rückwärts rechnend, die die [—]-Größen schrittweise bis zu den beiden in (28) auftretenden Größen $[p, 1, n]$ und $[q, 1, n]$ hin auffinden, so daß $S(n)$ als eine ganze rationale Funktion der beiden $\mathfrak{W}$-Größen p, q und der in (26) aufgestellten y-Größen erscheint. Letztere sind beständig gleich Eins, wenn der Index r gleich n oder größer ist; ferner sind für ein unter n gelegenes r die y gleich u oder gleich 1, je nachdem die Verbindungen (pp), (pq), (qp), (qq) als zählende Formen zu gelten haben oder nicht.

§ 164. Setzt man die Gleichungen (27) für $r = 1, 2, \cdots h$ an und fügt dazu die Gleichung (28), so lassen sich die zu $r = 1, 2, \cdots h - 1$ gehörenden [—]-Größen eliminieren, und man erhält für $S(n)$ eine lineare Verbindung aus $[p, h, n]$, $[q, h, n]$ von der Gestalt

$$S(n) = a(h, n)[p, h, n] + b(h, n)[q, h, n], \tag{30}$$

worin die Koeffizienten a und b ganze Funktionen der Größen p, q, y bedeuten. Ersetzt man ferner in (27) r durch h, so liefert die Verbindung

$$a(h, n) \times (27.\mathrm{a}) + b(h, n) \times (27.\mathrm{b})$$

auf der linken Seite wegen (30) den Ausdruck $S(n)$, d. h. es wird

$$\begin{aligned} S(n) = {} & [py(pp)_h a(h, n) + py(qp)_h b(h, n)] \cdot [p, h + 1, n] \\ & + [qy(pq)_h a(h, n) + qy(qq)_h b(h, n)] \cdot [q, h + 1, n]. \end{aligned}$$

Ersetzt man nun noch in (30) h durch $h + 1$ und vergleicht mit der vorstehenden Relation, so erhält man

$$\begin{aligned} a(h + 1, n) &= py(pp)_h a(h, n) + py(qp)_h b(h, n), \\ b(h + 1, n) &= qy(pq)_h a(h, n) + qy(qq)_h b(h, n). \end{aligned} \tag{31}$$

Aus diesen Rekursionsgleichungen ergeben sich die a, b durch Vorwärtsrechnen, sobald $a(1, n)$ und $b(1, n)$ bekannt ist. Für die letztgenannten Größen erhält man aber durch Vergleichung von (28) und (30)

$$a(1, n) = p, \quad b(1, n) = q. \tag{32}$$

Andrerseits folgt wegen (29) aus (30) für $h = n$

$$S(n) = a(n, n) + b(n, n). \tag{33}$$

Hiernach werden die Gleichungen (31) nur für die Nummern $h = 1$ bis $h = n - 1$ gebraucht. In diesem Falle hat aber die Größe $y(pp)_h$

für die vorkommenden Werte von h stets denselben Wert, und das Gleiche gilt von den anderen y-Größen. Man darf also jetzt bei den y-Größen den Index h fortlassen und hat dann, wenn zur Abkürzung noch

$$y(pp) = F, \quad y(pq) = G, \quad y(qp) = H, \quad y(qq) = K \tag{34}$$

geschrieben wird, nach den Gleichungen

$$a(1, n) = p, \quad b(1, h) = q,$$
$$a(h + 1, n) = pFa(h, n) + pHb(h, n),$$
$$b(h + 1, n) = qGa(h, n) + qKb(h, n),$$
$$S(n) = a(n, n) + b(n, n)$$

zu rechnen. Hierin ist n offenbar mindestens gleich 2 zu nehmen, da ja die Gruppen $G(2)$ erst bei zwei Zügen entstehen können. Schreibt man nun das vorstehende Gleichungssystem für zwei spezielle Werte von n — sagen wir für $n = k$ und $n = l$ — wirklich hin, so erkennt man, daß die Werte von $a(h, k)$ und $a(h, l)$ und ebenso die Werte von $b(h, k)$ und $b(h, l)$ übereinstimmen, solange h nicht über die kleinere der beiden Zahlen k und l hinausgeht. Man darf also in den obigen Rekursionsgleichungen, da h unterhalb n zu bleiben hat, für die Größen $a(h, n)$ und $b(h, n)$, auch $a(h, h)$ und $b(h, h)$ oder kürzer $a(h)$ und $b(h)$ schreiben, so daß die Gleichungen

$$\begin{aligned} a(h + 1) &= pFa(h) + pHb(h), \\ b(h + 1) &= qGa(h) + qKb(h) \end{aligned} \tag{35}$$

entstehen, denen man noch die Anfangsbedingungen

$$a(1) = p, \quad b(1) = q \tag{36}$$

nebst der Schlußgleichung

$$S(n) = a(n) + b(n) \tag{37}$$

hinzuzufügen hat.

Um die Lösung der aufgestellten Differenzengleichungen zu finden setzen wir mit der Veränderlichen t die Reihen

$$a = a(2)t^2 + a(3)t^3 + \cdots, \tag{38.a}$$
$$b = b(2)t^2 + b(3)t^3 + \cdots, \tag{38.b}$$
$$S = S(2)t^2 + S(3)t^3 + \cdots \tag{38.c}$$

an. Multipliziert man nun die Gleichungen (35) mit t^{h+1} und summiert nach h von $h = 1$ an, so wird unter Berücksichtigung von (36) und (38)

$$a = pFt(a + pt) + pHt(b + qt), \tag{39}$$
$$b = qGt(a + pt) + qKt(b + qt). \tag{40}$$

Setzt man zur Abkürzung noch

$$L = FK - GH,$$
$$A = t + pq(G + H - F - K)t^2,$$
$$B = 1 - (pF + qK)t + pqLt^2$$

und bildet die Verbindung

$$(qGt - qKt + 1) \times (39) + (pHt - pFt + 1) \times (40),$$

so erhält man daraus mit Unterdrückung der Zwischenrechnung

$$B(a + b + 1) = A. \tag{41}$$

Läßt man nun in der aus (22) folgenden Gleichung

$$\sum_n S(n) t^n = \sum_n \mathfrak{D}(u^x, n) t^n$$

die Summation nach n von $n = 1$ an laufen und legt dem bisher bedeutungslosen Zeichen $S(1) = \mathfrak{D}(u^x, 1)$ den Wert Eins bei, so nimmt die vorstehende Summe nach (38.c) den Wert $S + 1$ an. Andrerseits folgt aus (38) und (37) die Beziehung $S = a + b$. Damit geht (41) über in

$$\sum_n \mathfrak{D}(u^x, n) t^n = \sum_n \sum_x \mathfrak{U}(x, n) u^x t^n = A : B. \tag{42}$$

Hierdurch ist die Bestimmung der Ausdrücke $\mathfrak{D}(u^x, n)$ und $\mathfrak{U}(x, n)$ auf die Entwicklung des Quotienten $A : B$ zurückgeführt.

§ 165. Für die weitere Rechnung sind die fünf in § 162 aufgeführten Fälle zu trennen. Man erhält hierbei der Reihe nach folgende Werte von A und B.

I. $F = u, \quad G = H = K = 1, \quad L = u - 1,$
$A = t - pq(u - 1)t^2,$
$B = 1 - (pu + q)t + pq(u - 1)t^2;$

II. $F = H = K = 1, \quad G = u, \quad L = 1 - u,$
$A = t + pq(u - 1)t^2,$
$B = 1 - t - pq(u - 1)t^2;$

III. $F = K = u, \quad G = H = 1, \quad L = u^2 - 1,$
$A = t - 2pq(u - 1)t^2,$
$B = 1 - ut + pq(u^2 - 1)t^2;$

IV. $F = G = u, \quad H = K = 1, \quad L = 0,$
$A = t, \quad B = 1 - (pu + q)t,$

V. $F = H = u, \quad G = K = 1, \quad L = 0,$
$A = t, \quad B = 1 - (pu + q)t.$

Für die Fälle IV und V läßt sich der Koeffizient von t^n in der Entwicklung von $A:B$ sofort hinschreiben, und man erhält

$$\mathfrak{D}(u^x, n) = (pu + q)^{n-1}.$$

Dies ist aber nichts anderes, als die Grundgleichung für ein aus $n - 1$ Zügen zusammengesetztes *Bernoulli*sches Urnenschema: man darf sich also bei der weiteren Untersuchung auf die drei ersten Fälle beschränken.

Die Entwicklung des Quotienten $A:B$ kann für die drei ersten Fälle ohne eigentliche Schwierigkeiten recht wohl durchgeführt werden, indessen erscheinen hierbei die $\mathfrak{D}(u^x, n)$ zunächst in einer Gestalt, aus der sich nicht unmittelbar einfache Beziehungen ablesen lassen. Aus diesem Grunde sollen hier nur noch die Ausdrücke für die beiden niedrigsten Elemente $\mathfrak{D}(x)$ und $\mathrm{str}(x)$ abgeleitet werden, mit denen man bei den Anwendungen meistens ausreicht.

§ 166. Differentiiert man die Gleichung (42) wiederholt nach u und bezeichnet die Ableitungen von A und B durch Akzente, so erhält man, wenn nach der Differentiation $u = 1$ gesetzt wird, die Gleichungen

$$B^2 \sum \mathfrak{D}(x, n) t^n = BA' - AB',$$

$$B^3 \sum \mathfrak{D}(x^2 - x, n) t^n = B(BA'' - AB'') - 2B'(BA' - AB').$$

Hiermit stellt sich die Rechnung für die drei übrig gebliebenen Fälle folgendermaßen.

$$\text{I.}\quad A = t, \qquad A' = -pqt^2, \qquad A'' = 0,$$
$$B = 1 - t, \quad B' = -pt(1 - qt), \quad B'' = 0,$$
$$(1 - t)^2 \sum \mathfrak{D}(x, n) t^n = p^2 t^2,$$
$$(1 - t)^3 \sum \mathfrak{D}(x^2 - x, n) t^n = 2p^3 t^3 (1 - qt),$$

und daraus

$$(1 - t)^3 \sum \mathfrak{D}(x^2, n) t^n = p^2 t^2 (1 - t) + 2p^3 t^3 (1 - qt);$$

$$\text{II.}\quad A = t, \qquad A' = pqt^2, \qquad A'' = 0,$$
$$B = 1 - t, \quad B' = -pqt^2, \quad B'' = 0,$$
$$(1 - t^2) \sum \mathfrak{D}(x, n) t^n = pqt^2,$$
$$(1 - t)^3 \sum \mathfrak{D}(x^2 - x, n) t^n = 2p^2 q^2 t^4,$$

und daraus

$$(1 - t)^3 \sum \mathfrak{D}(x^2, n) t^n = pqt^2 (1 - t) + 2p^2 q^2 t^4;$$

$$\text{III.}\quad A = t, \qquad A' = -2pqt^2, \qquad A'' = 0,$$
$$B = 1 - t, \quad B' = -t(1 - 2pqt), \quad B'' = 2pqt^2,$$

$$(1-t)^2 \sum \mathfrak{D}(x, n) t^n = (1 - 2pq)t^2,$$
$$(1-t)^3 \sum \mathfrak{D}(x^2 - x, n) t^n = -2pqt^3(1-t) + 2(1-2pq)t^3(1-2pqt),$$

und daraus

$$(1-t)^3 \sum \mathfrak{D}(x^2, n) t^n = (1-2pq)t^2(1-t) - 2pqt^3(1-t)$$
$$+ 2(1-2pq)t^3(1-2pqt).$$

Ermittelt man hieraus die Werte von $\mathfrak{D}(x, n)$ und $\mathfrak{D}(x^2, n)$ und berechnet dann die Streuungen aus

$$\operatorname{str}(x, n)^2 = \mathfrak{D}(x^2, n) - \mathfrak{D}(x, n)^2,$$

so wird

$$\text{I.}\quad \mathfrak{D}(x, n) = (n-1)p^2,$$
$$\operatorname{str}(x, n)^2 = (n-1)p^2(1-p^2) + 2(n-2)p^3q,$$
$$\text{II.}\quad \mathfrak{D}(x, n) = (n-1)pq,$$
$$\operatorname{str}(x, n)^2 = (n-1)pq(1-pq) - 2(n-2)p^2q^2,$$
$$\text{III.}\quad \mathfrak{D}(x, n) = (n-1)(1-2pq),$$
$$\operatorname{str}(x, n)^2 = 2(n-1)pq(1-2pq) + 2(n-2)pq(p-q)^2.$$

Diese Ausdrücke sollen jetzt noch mit einem *Bernoulli*schen Schema verglichen werden, das $n - 1$ Züge umfaßt und durch die $\mathfrak{W}$-Größe

$$\mathfrak{W}(\text{weiß}) = P$$

charakterisiert ist. Bedeutet dann x die Menge der weißen Züge innerhalb eines Versuches, so ist für dieses Schema

$$\mathfrak{D}(x) = (n-1)P, \quad \operatorname{str}(x)^2 = (n-1)P(1-P).$$

Setzt man nun für P der Reihe nach die Werte

$$p^2 \text{ oder } pq \text{ oder } 1 - 2pq = p^2 + q^2,$$

so erhält man anstatt $\mathfrak{D}(x)$ die oben für $\mathfrak{D}(x, n)$ gegebenen Ausdrücke. Das Gruppenschema verhält sich also bezüglich des Argumentdurchschnitts gerade so, als wenn die $n - 1$ Gruppen $G(2)$ ebensoviele voneinander unabhängige Züge wären. Dagegen zeigen die Streuungen gegen das gewöhnliche Schema die Abweichungen

$$\text{I:}\quad + 2(n-2)p^3q,$$
$$\text{II:}\quad - 2(n-2)p^2q^2,$$
$$\text{III:}\quad + 2(n-2)pq(p-q)^2.$$

Im ersten und dritten Falle ist also die Streuung übernormal, im zweiten hingegen unternormal.

§ 167. Der vorstehend für die Gruppen $G(2)$ benutzte Weg ist auch noch für die Gruppen mit mehr als zwei Gliedern gangbar, wird aber mit wachsender Gliederzahl immer beschwerlicher. Deshalb will ich hier nur noch einige Resultate anführen.

Für die Gruppen $G(s)$ erhält man ähnlich wie vorher eine Darstellung von der Gestalt

$$\sum_n \mathfrak{D}(u^x, n) t^n = \sum_n \sum_x \mathfrak{U}(x, n) u^x t^n = A : B,$$

wo A und B ganze rationale Funktionen der Größen p, q, t und u bedeuten. Werden bei der Auszählung f zählende Formen berücksichtigt, so sind A und B nach u höchstens vom f-ten Grade. Für die Sequenzen von der Gestalt p^s erhält man eine Darstellung von der Form

$$\sum \mathfrak{D}(u^x, n) t^n = [A + B(u - 1)] : [C + D(u - 1)],$$

wo

$$A = t, \quad C = 1 - t,$$

$$(1 - pt)B = (1 - pt)(pt)^s - [1 - (pt)^{s-1}]pt^2,$$

$$(1 - pt)D = -(1 - pt)(pt)^s - [1 - (pt)^{s-1}]pt(1 - t)$$

ist. Hieraus folgt mit der Abkürzung

$$N = n - s + 1$$

für die beiden niedrigsten Elemente und für $n > 2s - 1$

$$\mathfrak{D}(x, n) = Np^s,$$

$$\operatorname{str}(x, n)^2 = Np^s(1 - p^s) + 2(N - 1)p^{s+1} + 2(N - 2)p^{s+2} + \cdots$$
$$+ 2(N - s + 1)p^{2s-1} - (s - 1)(2N - s)p^s \cdot p^s.$$

Ersetzt man hierin jedes angesetzte p^s durch $p^s + q^s$, so erhält man die Elemente für den Fall, daß die Formen p^s *und* q^s gezählt werden.

Das hier behandelte Gruppenschema ist mannigfacher Erweiterungen und Abänderungen fähig, und man kann sagen, daß eine halbwegs eingehende Diskussion der für einzelne Anwendungen in Betracht kommenden Fälle ein Buch für sich zu füllen vermag.

Neunzehnte Vorlesung.

Der Bayessche Satz.

§ 168. Als letztes Beispiel einer theoretischen Verteilung soll der als Umkehrung der *Bernoulli*schen Formel bezeichnete *Bayes*sche Satz dienen. Zu dem Ende schicken wir einige Vorbereitungen voraus.

Mit den positiven Größen a und b, die später große ganzzahlige Werte anzunehmen haben, bilde man den Ausdruck

$$F(x) = x^a (1 - x)^b. \tag{1}$$

Ferner werde eine Verteilungsfunktion $N(x)$ auf Grund der Bedingung konstruiert, daß $N(x)$ für echtgebrochene positive x bis auf einen konstanten Faktor C mit $F(x)$ übereinstimmen, sonst aber verschwinden soll. Der konstante Faktor C ist hierbei so zu bestimmen, daß das Integral von $N(x)$, genommen zwischen den Grenzen $\pm \infty$, den Wert Eins annimmt. Aus

$$N(x) = CF(x) = Cx^a (1 - x)^b \tag{2}$$

ergibt sich die zu der Verteilung $N(x)$ gehörige Summenfunktion $\mathfrak{S}(x)$ für echtgebrochene positive x vermittelst der Gleichung

$$\mathfrak{S}(x) = \int_0^x N(t)\, dt = C \int_0^x t^a (1 - t)^b dt; \tag{3}$$

ferner ist $\mathfrak{S}(x)$ beständig gleich 0 oder 1, je nachdem x unterhalb 0 oder oberhalb 1 liegt. Das vorstehende Integral ist uns bereits früher in § 38 (6) bei der Untersuchung eines Teilungsproblems begegnet; jetzt kommt indessen für uns die zugehörige Φ-Reihe in Betracht, die sich bei dieser Gelegenheit wieder als ein Hilfsmittel zur asymptotischen Darstellung von Funktionen großer Zahlen erweisen wird.

Für eine beliebige Funktion $T(x)$ ist der nach $N(x)$ genommene Durchschnitt aus

$$\mathfrak{D}(T) = \int_0^1 T(x) N(x)\, dx = C \int_0^1 T(x)\, x^a (1 - x)^b dx \tag{4}$$

zu berechnen. Setzt man für $T(x)$ die Potenz x^n, so wird

$$\mathfrak{D}(x^n) = C \int_0^1 x^{a+n} (1 - x)^b dx,$$

woraus nach § 30 (30)

$$\mathfrak{D}(x^n) = [C \Pi(a + n) \Pi(b)] : \Pi(a + b + n + 1)$$

folgt. Für $n = 0$ wird

$$1 = [C \Pi(a) \Pi(b)] : \Pi(a + b + 1), \tag{5}$$

so daß unter Elimination von C der Ausdruck

$$\mathfrak{D}(x^n) = [\Pi(a + n) \Pi(a + b + 1)] : [\Pi(a) \Pi(a + b + n + 1)]$$

entsteht. Setzt man zur Abkürzung

$$a + b + 2 = m,$$

$$a + 1 = mP, \quad b + 1 = mQ, \quad P + Q = 1,$$

so kann man schreiben

$$\mathfrak{D}(x^n) = \frac{mP}{m} \frac{mP + 1}{m + 1} \cdots \frac{mP + n - 1}{m + n - 1}. \tag{6}$$

Hieraus fließt für $n = 1$ und $n = 2$

$$\mathfrak{D}(x) = P, \qquad \mathfrak{D}(x^2) = P(mP + 1) : (m + 1), \tag{7}$$

so daß

$$\operatorname{str}(x)^2 = PQ : (m + 1) \tag{8}$$

wird. Damit sind die beiden ersten numerischen Elemente von $N(x)$ gefunden, ferner ergibt sich der Normalwert des Parameters h aus

$$2h^2PQ = m + 1. \tag{9}$$

Die Berechnung der übrigen Elemente läßt sich ebenfalls nach (6) ausführen. Jedoch ist dieser Weg etwas umständlich und deshalb ein anderer Rechnungsgang vorzuziehen.

§ 169. Mit den vorläufig beliebig angenommenen Parametern c, h werde angesetzt

$$u = h(x - c),$$

$$E = \exp(-2uw - w^2) = \sum_n \mathfrak{R}(u)_n (2w)^n,$$

$$Y = CEx^{a+1}(1 - x)^{b+1}.$$

Dann läßt sich die Ableitung von Y nach x zunächst in der Gestalt

$$dY : dx = EN(x)[-2hwx(1 - x) + (a + 1)(1 - x) - (b + 1)x]$$

schreiben. Führt man statt x das Hilfsargument u ein und setzt mit den vorhin aufgestellten Abkürzungen

$$m = a + b + 2, \qquad mP = a + 1, \qquad mQ = b + 1$$

die Ausdrücke

$$G = hm(P - c), \qquad H = -2h^2c(1 - c), \qquad K = 2h(2c - 1)$$

an, so erhält man für die Ableitung die Gleichung

$$h\,dY : dx = E[G + Hw - mu + Kwu + 2wu^2]N(x).$$

Integriert man die vorstehende Gleichung wieder nach x zwischen den Grenzen 0 und 1, so wird die linke Seite null, da Y an den Integrationsgrenzen verschwindet. Beachtet man ferner für die rechte Seite die Beziehung (4), so darf man schreiben

$$0 = (G + Hw)\mathfrak{D}(E) - (m - Kw)\mathfrak{D}(Eu) + 2w\mathfrak{D}(Eu^2). \tag{10}$$

Um hieraus die D-Koeffizienten zu finden, benutzen wir nach § 34 (16) die für die Polynome $\mathfrak{R}(u)_n$ geltende Gleichung

$$2u\mathfrak{R}_n = -(2n + 2)\mathfrak{R}_{n+1} - \mathfrak{R}_{n-1}, \tag{11}$$

aus der durch wiederholte Anwendung

$$4u^2\mathfrak{R}_n = 4(n + 1)(n + 2)\mathfrak{R}_{n+2} + (4n + 2)\mathfrak{R}_n + \mathfrak{R}_{n-2} \tag{12}$$

folgt. Diese Relationen können auch für negative n gebraucht werden, wenn man festsetzt, daß $\mathfrak{R}_n$ für ein negatives n stets null sein soll.

Summiert man die mit $(2w)^n$ multiplizierten Gleichungen (11) und (12) nach n von $-\infty$ bis $+\infty$ und bildet darauf den Durchschnitt nach $N(x)$, so erhält man unter Hinzufügung der Reihe für $\mathfrak{D}(E)$ die Entwicklungen

$$\mathfrak{D}(E) = \sum_n D_n (2w)^n,$$

$$2\mathfrak{D}(Eu) = -\sum_n (2n+2) D_{n+1}(2w)^n - \sum_n D_{n-1}(2w)^n,$$

$$4\mathfrak{D}(Eu^2) = 4\sum_n (n+1)(n+2) D_{n+2}(2w)^n + \sum_n (4n+2) D_n (2w)^n + \sum_n D_{n-2}(2w)^n.$$

Setzt man diese Reihen in (10) ein und spaltet darauf nach den Potenzen von w, so entsteht die Rekursionsformel

$$\begin{gathered} 4(n+1)(n+m)D_{n+1} = \\ (2Kn - 4G)D_n - 2(2n - 1 + m + H)D_{n-1} + KD_{n-2} - D_{n-3}, \end{gathered} \tag{13}$$

aus der sich, da $D_0 = 1$ gegeben ist, sämtliche D-Koeffizienten der Reihe nach finden lassen.

Setzt man für die Größen c, h zunächst die Normalwerte ein, so ist nach (7) und (9)

$$c = P, \quad 2h^2 PQ = m + 1,$$

und man erhält

$$G = 0, \quad H = -m - 1, \quad K = 2h(P - Q),$$

woraus

$$\begin{gathered} 4(n+1)(n+m)D_{n+1} = \\ 4hn(P-Q)D_n - (4n-4)D_{n-1} + 2h(P-Q)D_{n-2} - D_{n-3} \end{gathered} \tag{14}$$

folgt. Setzt man hierin n der Reihe nach gleich $0, 1, \ldots$, so wird zunächst, wie es sein muß, $D_1 = 0$ und $D_2 = 0$, ferner

$$4(m+2)D_3 = h(P-Q), \tag{15.a}$$

$$16(m+2)(m+3)D_4 = 2h^2(1 - 5PQ) - 1, \tag{15.b}$$

$$\cdots\cdots$$

Dividiert man die Rekursionsformel (14) durch m und läßt die Zahlen a, b und folgeweise auch m über alle Grenzen wachsen, so gehen auf der rechten Seite von (14) die Koeffizienten der D-Größen gegen Null. Folglich konvergieren dann alle D-Koeffizienten, von D_0 abgesehen, gegen Null, so daß sich die Φ-Reihe auf den Anfangsterm $\Phi(u)$ reduziert.

Will man die *Bayes*sche Formel in derjenigen Gestalt erhalten, in der sie gewöhnlich gegeben wird, so hat man statt der normalen c, h etwas andere Werte anzusetzen. Zunächst sei zur Abkürzung

$$a + b = e, \quad p = a : e, \quad q = b : e, \tag{16}$$

so daß $m = e + 2$ wird. Weiter sollen die Parameter c, h durch die Gleichungen

$$c = p, \quad 2h^2 pq = e \tag{17}$$

bestimmt werden. Dann wird

$$G = h(1 - 2p), \quad H = -e, \quad K = 2h(2p - 1),$$

$$2Kn - 4G = 4h(p - q)(n + 1), \quad 2n - 1 + m + H = 2n + 1,$$

womit die Rekursionsformel die Gestalt

$$\begin{gathered} 4(n + 1)(n + 2 + e)D_{n+1} = \\ h(4n + 4)(p - q)D_n - (4n + 2)D_{n-1} + 2h(p - q)D_{n-2} - D_{n-3} \end{gathered} \tag{18}$$

annimmt. Läßt man nach Division mit e die Zahlen a, b wieder unbegrenzt wachsen, so ergibt sich wie vorhin, daß die Koeffizienten $D_1, D_2, \ldots$ gegen Null gehen, daß sich also die Φ-Reihe auf ihr Anfangsglied reduziert. Setzt man ferner für n der Reihe nach $0, 1, \ldots$, so erhält man nach einigen Reduktionen

$$(2 + e)D_1 = h(p - q), \tag{19.a}$$

$$(2 + e)(3 + e)D_2 = h^2(2 - 11pq) - 3, \tag{19.b}$$

$$\cdots\cdots$$

Der Parameter $c = p$ entspricht dem dichtesten Werte der Verteilungsfunktion $N(x)$. Differentiiert man nämlich $N(x)$, um das Maximum von N zu finden, logarithmisch, so erhält man die Bedingung

$$\frac{a}{x} - \frac{b}{1 - x} = 0,$$

die in der Tat durch $x = p$, $1 - x = q$ befriedigt wird.

Beschränkt man sich unter Voraussetzung sehr großer Werte von a und b auf das Anfangsglied der Φ-Reihe, so erhält man für die Summenfunktion $\mathfrak{S}(x)$ die nachstehenden von den Zahlen a, b abhängenden Beziehungen

$$a + b = e, \quad p = a : e, \quad q = b : e, \tag{20}$$

$$2h^2 pq = e, \quad u = h(x - p), \tag{21}$$

$$2\mathfrak{S}(x) - 1 = \Phi(u), \tag{22}$$

welche bei passender Deutung den jetzt zu besprechenden Satz von *Bayes* liefern.

§ 170. Als Ausgangspunkt dient folgende Aufgabe. Aus einer Urne, die weiße und schwarze Kugeln in unbekannter Menge enthält, sind unter jedesmaliger Zurücklegung der gezogenen Kugel a weiße und b schwarze Züge erfolgt; gesucht wird die W. des einzelnen weißen Zuges, die wir mit w bezeichnen wollen.

Die herkömmliche Behandlung der Aufgabe stützt sich auf das

*Bayes*sche Prinzip, wobei im besondern die Formeln (10) und (11) aus § 21 in Betracht kommen. Danach sind zunächst die fünf Funktionen

$$F(x), \quad \mathfrak{W}(x)_v, \quad \mathfrak{W}(x)_n, \quad V(x), \quad N(x)$$

zu unterscheiden, die folgende Bedeutung haben. Die erste Funktion $F(x)$ drückt die $\mathfrak{W}$. aus, mit der man das beobachtete Ereignis zu erwarten hat, wenn die Füllungsgröße w den Wert x besitzt. Zweitens bedeuten $\mathfrak{W}(x)_v$ und $\mathfrak{W}(x)_n$ die „vorläufige" und die „nachträgliche" $\mathfrak{W}$. der *Annahme*, daß w unterhalb x liege. Drittens endlich bedeuten $V(x)$ und $N(x)$ die nach x genommenen Ableitungen von $\mathfrak{W}(x)_v$ und $\mathfrak{W}(x)_n$.

Da die möglichen Werte von w in dem Intervall von $w = 0$ bis $w = 1$ enthalten sind, so müssen die beiden Funktionen $\mathfrak{W}_v$ und $\mathfrak{W}_n$ ihrer Bedeutung wegen für negative x beständig null sein. Ferner gehen beide Funktionen, wenn x von 0 bis 1 läuft, wachsend oder doch wenigstens nicht abnehmend gegen Eins und behalten nachher diesen Wert für die oberhalb Eins gelegenen x beständig bei. Die $\mathfrak{W}(x)$ verhalten sich also genau so, wie die Summenfunktionen eines stetigen K.-G., dessen mögliche Argumentwerte auf das Intervall von $x = 0$ bis $x = 1$ beschränkt sind. Die zugehörigen Verteilungsfunktionen sind dann durch die Ableitungen $V(x)$ und $N(x)$ gegeben, die nach dem Gesagten außerhalb des Intervalles von 0 bis 1 beständig verschwinden.

Bei der Aufsuchung von w wollen wir zunächst an die Umstände anknüpfen, die beim wirklichen Rechnen eintreten. Um die Vorstellung zu fixieren, möge angenommen werden, daß das gesuchte w abgerundet mit drei Dezimalen anzusetzen sei, daß es also mit einem Gliede der Zahlenreihe

$$0.000, \quad 0.001, \quad 0.002, \quad \ldots \quad 0.999, \quad 1.000 \tag{23}$$

zusammenfalle. In dieser Reihe vertritt wegen der Abrundung jedes Glied nicht nur die hingeschriebene Zahlengröße, sondern auch noch die benachbarten Werte bis zu den nächstgelegenen Wechselpunkten hin. Infolgedessen ist die nachträgliche $\mathfrak{W}$. einer Annahme — sagen wir z. B. der Annahme $w = 0.368$ — durch die Differenz

$$\mathfrak{W}(0.3685)_n - \mathfrak{W}(0.3675)_n$$

gegeben. Berechnet man in dieser Weise zu allen in (23) vorkommenden Annahmen die zugehörigen $\mathfrak{W}$-Differenzen, so entsteht eine Tabelle, die offenbar nichts anderes ist, als die zu den Argumenten (23) aus der Summenfunktion $\mathfrak{W}(x)_n$ berechnete Verteilungstafel. Da der Gang der Verteilungswerte im wesentlichen den Verlauf der Funktion $N(x)$ wiederspiegeln muß, so wird in der Tafel ein Argument vorkommen, dessen $\mathfrak{W}$. die übrigen $\mathfrak{W}$-Größen übertrifft. Dieses Argument mit größter $\mathfrak{W}$. wird nun als sogenannte „wahr-

scheinlichste“ Annahme für das gesuchte w genommen, indem man von der Auffassung ausgeht, daß von je zwei Annahmen mit ungleicher $\mathfrak{W}$. die wahrscheinlichere den Vorzug verdiene.

§ 171. Die vorstehende Behandlung der betrachteten Aufgabe enthält wegen der vorläufig in das Belieben des Rechners gestellten Abrundung noch einen willkürlichen Bestandteil und zugleich eine Unschärfe. Um diesen Mangel zu beseitigen, denken wir uns jetzt die Abrundung schrittweise verkleinert und schließlich unendlich klein gemacht. Bedeutet in dem Falle der unendlich kleinen Abrundung $2T$ die unendlich kleine konstante Teilstreckenlänge und x die Mitte einer Teilstrecke, so ist die Verteilungstafel der $\mathfrak{W}$-Größen mit den Differenzen

$$\mathfrak{W}(x+T)_n - \mathfrak{W}(x-T)_n = 2T \cdot N(x) \tag{24}$$

zu berechnen. Da diese $\mathfrak{W}$-Größen den Werten von $N(x)$ proportional sind, so ist jetzt die wahrscheinlichste Lösung aus der Bedingung zu ermitteln, daß $N(x)$ ein Maximum wird.

Um das Maximum von $N(x)$ zu finden, muß man den Ausdruck dieser Funktion kennen. Nun stellt das *Bayes*sche Prinzip zwischen $F(x)$, $V(x)$ und $N(x)$ die Beziehung

$$N(x) = CF(x)V(x)$$

her, in der die Konstante C aus dem Umstande zu bestimmen ist, daß das Integral über $N(x)$, zwischen den Grenzen $\pm\infty$ genommen, den Wert Eins besitzt. Der Faktor $F(x)$ ist bei der vorliegenden Aufgabe durch

$$F(x) = x^a(1-x)^b$$

gegeben, denn der betrachtete Urnenversuch besteht aus a weißen und b schwarzen Zügen, die in bestimmter Reihenfolge beobachtet worden sind. Der Faktor $V(x)$ wird herkömmlicher Weise konstant gesetzt, mit der Motivierung, daß man sich vor Anstellung des Versuches über den Wert von w in völliger Ungewißheit befinde und zwischen den verschiedenen möglichen Annahmen hinsichtlich ihrer vorläufigen $\mathfrak{W}$. keinen Unterschied machen könne. Ein solcher Ansatz ist allerdings, wie bereits in § 24 erwähnt wurde, häufig willkürlich und nicht einmal aus dem Prinzip des mangelnden Grundes ausreichend zu rechtfertigen; überdies kann es vorkommen, daß durch die Einführung eines konstanten $V(x)$ wesentliche Umstände, die bereits vor Anstellung des Versuchs bekannt gewesen sind, tatsächlich beiseite geschoben werden. Wir wollen uns indessen mit diesen Einwendungen nicht aufhalten, sondern bei der üblichen Bestimmung bleiben, da es uns hier auf den *Bayes*schen Satz in seiner hergebrachten Form ankommt. Setzt man demgemäß die Funktion $V(x)$ in dem Gebiete

der möglichen x konstant, so erscheint $N(x)$ innerhalb dieses Gebietes in der Gestalt

$$N(x) = EF(x) = Ex^a(1-x)^b,$$

wo der Faktor E eine Konstante bedeutet. Da ferner im Gebiete der außermöglichen x die Funktion $N(x)$ beständig verschwindet, so sind wir bei dem Ausdrucke angelangt, der oben in §§ 168—169 als Ausgangspunkt genommen wurde; die dort gefundenen Ergebnisse sind also jetzt unmittelbar zu verwenden.

Für die wahrscheinlichste Lösung der betrachteten Urnenaufgabe ergibt sich nunmehr der Wert

$$w = p = a:(a+b).$$

Diese Bestimmung gilt ohne Rücksicht darauf, ob die Zügezahlen a, b groß sind oder nicht. Weiter ergeben sich, wenn a und b hinreichend groß sind, aus (20)—(22) die Gleichungen

$$2h^2ab = (a+b)^3, \qquad u = h(x-p),$$

$$2\,\mathfrak{W}(x)_n - 1 = \Phi(u).$$

Vorstehende Gleichung wird benutzt, um die „wahrscheinlichen" Grenzen für die gesuchte Unbekannte aufzustellen. Ist nämlich g eine positive Größe, so ergibt sich für die $\mathfrak{W}$. die Annahme, daß w zwischen den Grenzen $p \pm g$ liege, der Ausdruck

$$\mathfrak{W}(p+g)_n - \mathfrak{W}(p-g)_n = \Phi(hg).$$

Bestimmt man nun g aus der Bedingung $\Phi(hg) = 0.5$ oder $hg = 0.477$, so kann man Eins gegen Eins wetten, daß w zwischen den Grenzen $p \pm g$ liege. Diese wahrscheinlichen Grenzen dienen offenbar demselben Zwecke, wie die von uns eingeführten Streuungsgrößen, nämlich zur Abschätzung des Vertrauens, das der gefundenen Lösung zu schenken ist.

Ist die Zügezahl $a + b$ groß, so erwächst daraus neben der Erhöhung des der Lösung zu schenkenden Vertrauens noch ein anderer Vorteil, der jetzt zu besprechen ist. Man denke sich zu x als Abszisse die Werte der Funktionen $F(x)$, $V(x)$ und $N(x)$ als Ordinaten abgetragen. Dann ist der Verlauf der Kurve F vollständig bestimmt, während der Verlauf von V ungewiß ist, weil sich im allgemeinen jede bestimmte Festsetzung darüber anfechten läßt. Indessen wird man der Regel nach über die Kurve V das Eine aussagen dürfen, daß sie im allgemeinen stetig verlaufen und keine „steilen Berge" besitzen wird, d. h. keine Stellen, wo die Ordinate rasch zu einem starken Maximum aufsteigt und ähnlich rasch wieder abfällt. Denn ein solcher Berg würde bedeuten, daß für die Annahmen, die den x in der nächsten Umgebung des Maximums entsprechen, schon von vornherein eine besonders große $\mathfrak{W}$. vorhanden sei. Die Kurve N ent-

steht dann durch die Multiplikation der Ordinaten V mit den Ordinaten F und einem konstanten Faktor. Kann man nun den anzustellenden Versuch so einrichten, daß die Kurve F einen sehr steilen Berg besitzt, im übrigen aber ganz nahe an der Abszissenachse verläuft, so wird sie diese charakteristische Eigenschaft unter der über V gemachten Voraussetzung auf die Kurve N übertragen; es wird mit anderen Worten der Spielraum, der für die Einzeichnung der Kurve V besteht, in dem Verlaufe von N nur stark abgeschwächt zum Vorschein kommen. Die besondere Form von F bewirkt also, daß sich bei dem Übergange von den vorläufigen auf die nachträglichen WW. die den ersteren anhaftende Ungewißheit stark vermindert. Bei der vorliegenden Urnenaufgabe ist nun, wenn $a+b$ sehr groß genommen wird, der verlangte steile Berg vorhanden, wie sich aus dem Verhalten des mit einem konstanten V abgeleiteten $N(x)$ ergibt.

§ 172. Nachdem wir vorstehend den *Bayes*schen Satz in seiner üblichen Gestalt kennen gelernt haben, wollen wir zusehen, wie weit er einer kritischen Prüfung gegenüber Stand hält.

Die W. der Annahme $w = x$ war, wenn man eine scharfe Bestimmung, also eine unendlich kleine Abrundung voraussetzt, nach (24) durch den Differentialausdruck $2T \cdot N(x)$ gegeben. Das führt zu der Frage, welcher Wert denn irgend einer Aussage, sei es auch der „wahrscheinlichsten", beizumessen sei, wenn ihre W. unendlich klein ist — man pflegt doch z. B. bei *Ereignissen* die Fälle mit unendlich kleiner W. als etwas belangloses zu behandeln. Das aufgeworfene Bedenken wird auch nicht durch den Umstand gehoben, daß die unendlich kleinen W-Größen unvermeidlich bei jeder irgendwie gearteten Lösung der vorgelegten Urnenaufgabe auftreten, sobald die möglichen Werte des gesuchten w in unendlicher Anzahl vorhanden sind. Vielmehr führt gerade dieser Umstand zu der weiteren Frage, ob denn nicht etwa der Ansatz der Lösung von vornherein auf einen falschen Weg geraten sei. Der gleiche Zweifel stellt sich ein, wenn man den anderen bedenklichen Punkt in dem *Bayes*schen Satze, nämlich die Bestimmung von $V(x)$ prüft. Denn neben den Beispielen, bei denen die Konstanz von $V(x)$ einwandfrei aus dem Prinzip des mangelnden Grundes folgt, gibt es genug andere Aufgaben, bei denen sich ganz bestimmte Gründe *gegen* jene Konstanz geltend machen lassen; man denke z. B. an den Fall eines Würfels mit den Kantenlängen 2, 3, 4. Daß in solchen Fällen die Ansetzung irgend eines *bestimmten* veränderlichen $V(x)$ ebenfalls angefochten werden kann, ist offenbar keine Stütze für das konstante $V(x)$, sondern nur ein Grund mehr für die Auffassung, daß der ganze Lösungsansatz verfehlt sei. In der Tat beruht die Beseitigung der angeführten Schwierigkeiten wesentlich darauf, daß bei der Lösung der gestellten Aufgabe die Heranziehung des *Bayes*schen Prinzips in Wahrheit überflüssig ist, weil der gewollte Zweck weit

kürzer und einfacher auf einem anderen Wege erreicht werden kann. Wie die Dinge zur Zeit liegen, kommt dem genannten Prinzip in keiner Weise die Bedeutung zu, die ihm in den meisten Darstellungen der W.-R. aus alter Gewöhnung beigelegt wird; man könnte es aus den Lehrbüchern ohne Schaden ausmerzen, wenn man es nicht zum Verständnis mancher älterer Arbeiten nötig hätte. Um das zu zeigen, sollen zunächst einige Bemerkungen allgemeiner Natur vorausgeschickt werden.

§ 173. Der Sprachgebrauch umfaßt mit dem Worte „Beobachtung", soweit es sich dabei um ziffernmäßig bestimmbare Dinge handelt, zwei große Gebiete empirischer Wahrnehmung, nämlich einerseits die *messenden* „physikalischen", andrerseits die *zählenden* „statistischen" Methoden. Die eigentliche Messung kommt in letzter Linie darauf hinaus, daß ausgedehnte Dinge, wie z. B. Winkel, Längen, Massen usw. mit gleichartigen Maßeinheiten verglichen werden, während die statistische Methode zunächst nur abzählt, wie oft in einer Reihe von Fällen ein gewisses Ereignis eingetreten ist. Man kann nun fragen, ob hierin ein Anlaß liege, gemessene und gezählte Größen *grundsätzlich* verschieden zu behandeln, wie das seither Regel gewesen ist.

Wenn für eine Unbekannte X, z. B. für einen Winkel, durch einmalige Messung der Wert B gefunden worden ist, so fällt es verständigerweise niemand ein, erst tiefsinnige Betrachtungen darüber anzustellen, ob etwa die Annahme $X = B$ oder eine andere die größte Wahrscheinlichkeit besitze. Man nimmt vielmehr kurzweg für das gesuchte X das beobachtete B, und zwar aus dem ebenso einfachen wie triftigen Grunde, weil kein anderer Wert von X gemessen worden ist, und im vorliegenden Falle jeder von $X = B$ verschiedene Ansatz auf pure Willkür hinauslaufen würde. Das geschieht, trotzdem man erfahrungsgemäß weiß, daß das gefundene B mit einen „Beobachtungsfehler" von unbekanntem Betrage behaftet ist. Die Messung erscheint hierbei als die Antwort auf eine an die Natur gestellte Frage und darf ohne *zwingenden* Grund nicht abgeändert werden — die ungerechtfertigte Abänderung gilt als Fälschung. Denn die *Natur* begeht bei der Antwort, die sie gibt, keinen Fehler, wohl aber *wir* bei der Deutung der Antwort; *darum ist der sogenannte Beobachtungsfehler in Wahrheit gar kein Fehler der Beobachtung, sondern nur der Deutung, die wir der Beobachtung unterlegen.*

Faßt man diese Bemerkungen kurz zusammen, so liegt der wesentliche Punkt offenbar darin, daß man sich bei dem Ansatz $X = B$ lediglich an die Tatsache der Beobachtung B hält. Das gerade Widerspiel hierzu ist nun die Behandlung der statistischen Beobachtung, die wir bei dem *Bayes*schen Satze kennen gelernt haben. Denn die Lösung der vorgelegten Urnenaufgabe wird an das Maximum des Ausdrucks

$$N(x) = CF(x)\,V(x)$$

gebunden, also wegen des Faktors $V(x)$ tatsächlich von dem abhängig gemacht, was man vor Anstellung der Beobachtung *nicht* weiß oder auch unter Umständen nicht wissen *will*, während sonst überall der Grundsatz gilt, daß eine Aussage sich auf das stützen solle, was man weiß. Es ist darum auch nicht weiter zu verwundern, daß der *Bayes*sche Satz bei kritischer Prüfung auf Schwierigkeiten führt. Um diesen Schwierigkeiten zu entgehen lassen wir jetzt das *Bayes*sche Prinzip gänzlich beiseite und kleiden das Problem zunächst in folgende zwei Fragen: 1) was ist der wahre Inhalt der eigentlichen Beobachtung, 2) was ist das vorläufig unbekannte Objekt der Beobachtung? Es wird sich zeigen, daß die Urnenaufgabe in der Hauptsache schon durch die bloße Form der Fragestellung ihre Lösung findet.

§ 174. Der statistische Versuch bildet seinem Wesen nach eine Kollektivreihe und ist deshalb auch als solche zu behandeln. Bei dem betrachteten Urnenversuche bestehen die Glieder der Kollektivreihe aus den $a + b$ Zügen, ferner ist das als Argument auftretende veränderliche Merkmal der einzelnen Glieder durch die Farbe der gezogenen Kugeln oder, arithmetisch ausgedrückt, durch die rH. gegeben, mit der die weißen Kugeln bei jedem Zuge vorkommen. Das Argument x besitzt also die beiden möglichen Werte $x = 0$ und $x = 1$. Ordnet man nun die Urliste der $a + b$ Züge in die Verteilungstafel um, so erhält man eine „beobachtete“ Verteilung $\mathfrak{U}(x)$ mit den beiden Ordinaten

$$\mathfrak{U}(0) = b : (a + b), \qquad \mathfrak{U}(1) = a : (a + b).$$

Damit ist der Inhalt der Beobachtung genau beschrieben. Zugleich ergibt sich aber daraus auch das Objekt der Beobachtung, denn wer $\mathfrak{U}(x)$ beobachtet, der sucht die „theoretische“ Verteilung $\mathfrak{W}(x)$, die man bei einer gleichmäßigen Erschöpfung der gleichmöglichen Fälle oder, kürzer ausgedrückt, bei einer vollständigen Ausgleichung des Zufalls beobachten würde. Die Auffindung von $\mathfrak{W}(x)$ ist überdies äquivalent mit der Bestimmung der Unbekannten w des *Bayes*schen Satzes, denn w ist der Argumentdurchschnitt von $\mathfrak{W}(x)$, und umgekehrt erhält man aus w sofort auch $\mathfrak{W}(x)$.

Die Aufgabe stellt sich also jetzt so: gesucht wird das Ordinatenpaar $\mathfrak{W}(x)$, beobachtet ist statt dessen das Ordinatenpaar $\mathfrak{U}(x)$, das bei vollständiger Ausgleichung des Zufalls mit $\mathfrak{W}(x)$ zusammenfallen würde, in Wirklichkeit jedoch wegen der unvermeidlichen Einwirkungen des Zufalls von $\mathfrak{W}(x)$ um unbekannte Beträge abweicht. Bei einer solchen Fassung der Aufgabe ist nun aber schlechterdings kein Grund zu erkennen, weshalb man die vorgelegte statistische Beobachtung anders behandeln müßte, als den oben besprochenen Fall einer einzelnen Messung. Mit anderen Worten: man setzt $\mathfrak{U}(x)$ für $\mathfrak{W}(x)$, ohne erst nach der größeren oder geringeren, aber stets un-

endlich kleinen, $\mathfrak{W}$. dieses Ansatzes zu fragen und ohne sich um das zu kümmern, was man vor Anstellung der Beobachtung *nicht* weiß. Ausschlag gebend ist, daß für $\mathfrak{W}(x)$ dies eine und nur dies eine $\mathfrak{U}(x)$ beobachtet wurde — damit ist die Sache erledigt, für den statistischen Versuch so gut wie für eine einmalige Messung. *Daß dadurch die Heranziehung des Bayesschen Prinzips zu einem völlig überflüssigen Ballast wird, ist einleuchtend, aber sicherlich kein Nachteil und noch weniger ein Hindernis, eine einfache Aufgabe auf einfache Weise zu behandeln.*

Der *Bayes*sche Satz stellt, wie wir gesehen haben, außer der wahrscheinlichsten Lösung auch noch ein Unsicherheitsmaß in Gestalt der wahrscheinlichen Grenzen der Unbekannten w auf. Es ist indessen nicht nötig, hierauf einzugehen, weil wir den Gegenstand bereits in der XV. Vorlesung in sehr viel allgemeinerer Form behandelt haben.

In den vorstehenden Erörterungen war der Nachdruck darauf gelegt worden, daß kein Anlaß vorliege, Messungen und Zählungen grundsätzlich verschieden zu behandeln. Im übrigen sind natürlich zwischen den beiden Klassen von Beobachtungen Unterschiede vorhanden. In dieser Beziehung möge hier nur auf einen Punkt hingewiesen werden. Der statistische Versuch bietet, weil er seinem Wesen nach Kollektivreihe ist, unter gewissen Umständen zugleich auch die Möglichkeit, Streuungsgrößen und damit Unsicherheitsmaße zu berechnen. Bei der einmaligen Messung einer Unbekannten fällt diese Möglichkeit fort.

§ 175. Zum Schlusse wollen wir die vorstehenden Bemerkungen über die statistische Beobachtung noch von den Einschränkungen befreien, die aus der Besonderheit des zugrunde gelegten Urnenschemas entspringen. Die statistische Zählung liefert, allgemein gesprochen, als Ergebnis eine gewisse beobachtete Verteilung $\mathfrak{U}$, die von einem oder von mehreren Argumenten abhängt. Gesucht wird dazu die entsprechende theoretische Verteilung $\mathfrak{W}$, die außer den Argumenten von $\mathfrak{U}$ auch noch gewisse Parameter enthält, die als die eigentlichen Unbekannten des Problems auftreten. Der Ansatz zur Lösung der Aufgabe besteht dann darin, daß man für die beobachteten Wertsysteme der Argumente die Gleichungen $\mathfrak{W} = \mathfrak{U}$ aufstellt, deren weitere Behandlung sich verschieden gestaltet, je nachdem einer der nachstehenden drei Fälle vorliegt.

Der erste Fall tritt ein, wenn die Gleichungen $\mathfrak{W} = \mathfrak{U}$ durch ein gewisses Wertsystem der Unbekannten befriedigt werden und gleichzeitig zur völligen Bestimmung der Unbekannten ausreichen. Die Aufgabe ist dann, wie man zu sagen pflegt, *bestimmt*; ihre Lösung ist gefunden, sobald man das System $\mathfrak{W} = \mathfrak{U}$ nach den Unbekannten aufgelöst hat.

Wenn zweitens die Gleichungen $\mathfrak{W} = \mathfrak{U}$ durch unendlich viele Wertsysteme der Unbekannten befriedigt werden, so ist die Aufgabe

schlechthin unbestimmt. Man mag dann immerhin versuchen, an der Hand des *Bayes*schen Prinzips wenigstens eine wahrscheinlichste Lösung aufzustellen, muß sich aber dabei gegenwärtig halten, daß keine logische Deduktion die tatsächliche Lücke in dem Beobachtungsmaterial wirklich auszufüllen vermag. Aus einem Nichtwissen läßt sich kein Wissen erzeugen — über diesen Satz hilft keine Kunst der Wahrscheinlichkeitsrechnung hinweg.

Drittens endlich kann es vorkommen, daß die Gleichungen $\mathfrak{W} = \mathfrak{U}$ durch kein einziges Wertsystem der Unbekannten befriedigt werden, sondern einander widersprechen, weil den beobachteten $\mathfrak{U}$-Größen die unausgeglichenen Wirkungen des Zufalls anhaften. Dieser Fall ist besonders wichtig, zumal man ihn sehr oft absichtlich herbeiführt. Wir werden ihn jedoch hier nicht weiter behandeln, weil der Gegenstand in die sogenannte Ausgleichungsrechnung gehört, die ein besonderes Kapitel der angewandten Mathematik mit eigenen Problemen und Methoden bildet. Nur beiläufig mag bemerkt werden, daß man gerade in der Ausgleichungsrechnung eine Zeitlang ausgiebig mit dem *Bayes*schen Prinzip und seinen wahrscheinlichsten Lösungen operiert hat, daß jedoch diese Behandlungsweise in dem Augenblicke antiquiert war, als die „Theoria combinationis etc.“ von *Gauß* erschien. Freilich ist das auch heute noch nicht allgemein zum Bewußtsein gekommen.

Der Ansatz $\mathfrak{W} = \mathfrak{U}$, der jedes beobachtete $\mathfrak{U}$ zu seinem Rechte kommen läßt, bildet unbeschadet praktisch erlaubter Abkürzungen und Vereinfachungen das Fundament für die rationelle Behandlung statistischer Beobachtungen, sobald es sich darum handelt, aus dem beobachteten $\mathfrak{U}(x)$ eine theoretische Verteilung $\mathfrak{W}(x)$ zu ermitteln, deren Gestalt je nach den Umständen hypothetisch oder auf Grund anderweiter Erfahrungen gegeben sein kann. Genau die gleiche Form des Ansatzes kommt zum Vorschein, wenn man daran geht, statt einer einzelnen Messung *Reihen* von Messungen zu untersuchen, so daß auch bei der verallgemeinerten Aufgabe kein Grund besteht, Zählung und Messung *grundsätzlich* voneinander zu scheiden.

Zwanzigste Vorlesung.

Numerische Bearbeitung: direkte Mittelbildung.

§ 176. Die Frage, wie die rechnerische Bearbeitung einer beobachteten Kollektivreihe zu gestalten sei, war bisher nur flüchtig gestreift worden. Wir haben daher diesen Gegenstand, der zugleich den Abschluß unserer Darstellung bilden soll, noch des näheren zu erörtern.

Da die Dinge, die Objekt der Kollektivmaßlehre werden können, eine äußerst bunte Mannigfaltigkeit bilden, so wird man von vornherein darauf verzichten, für den Rechnungsgang eine feste, auf alle vorkommenden Fälle gleich gut passende Schablone herrichten zu wollen, zumal bei der Behandlung beobachteter Zahlen überall auch die Beschaffenheit der benutzten Beobachtungen mitzusprechen hat. Der Rechner wird sich ja stets die Entstehungsweise der vorgelegten Zahlen gegenwärtig zu halten und dabei zu fragen haben, ob und wieweit die etwa vorhandenen besonderen Umstände einen besonderen, von der gewöhnlichen Regel abweichenden Weg notwendig machen; das Werk von *Fechner* enthält in dieser Beziehung manche für die Kollektivmaßlehre beherzigenswerte Winke und Beispiele. Unter diesen Umständen werden wir uns also hier darauf beschränken, allgemeine Gesichtspunkte hervorzuheben und an Beispielen zu erläutern.

Wählt man als mathematische Grundlage der Kollektivmaßlehre die Φ-Reihe, so erscheint als nächstes Ziel der Rechnung die Ermittelung der Größen, die wir unter dem Namen „numerische Elemente" zusammengefaßt haben. Hierbei entsteht sogleich die Frage, wieviele von diesen Elementen zu berechnen oder mit anderen Worten, wieviele Glieder der Φ-Reihe mitzunehmen seien. Denn wenn man auf der einen Seite wünschen muß, daß die Rechnung aus einer vorgelegten Kollektivreihe alles das heraushole, was die Beobachtung an erkennbarer Gesetzmäßigkeit in sich schließt, so wird man auf der anderen Seite, um Rechenarbeit zu sparen, mit der Gliederanzahl der Φ-Reihe nicht weiter gehen wollen, als jedesmal gerade nötig ist. Unter solchen Umständen besteht der natürlichste und seither auch vorwiegend innegehaltene Weg darin, daß man schrittweise vorgeht und in dem Bilde der beobachteten Kollektivreihen zunächst nur die gröberen, durch die Elemente niedrigerer Ordnung charakterisierten Züge aufsucht, mit dem Vorbehalt, die Fragestellung nach und nach auch auf die weniger augenfälligen Züge auszudehnen. Eine gute Erläuterung hierzu liefert das Verfahren der Meteorologie, bei dem wir deshalb einen Augenblick verweilen wollen.

§ 177. Die Meteorologie ist in demjenigen Teile, den man füglich als den „beschreibenden" bezeichnen kann, der Hauptsache nach eine — allerdings rohe — Kollektivmaßlehre der Witterungserscheinungen. Um das deutlich zu machen, ist nur nötig, das Verfahren des Meteorologen in die uns geläufige Ausdrucksweise der Kollektivmaßlehre zu übersetzen, wobei wir der Kürze halber und um die Vorstellung zu fixieren, ein bestimmtes meteorologisches Element, z. B. die Lufttemperatur, herausgreifen wollen.

Die thermometrischen Aufzeichnungen des meteorologischen Beobachtungstagebuchs bilden eine Urliste, deren Glieder zunächst durch ihre zeitliche Reihenfolge unterschieden sind. Von den veränderlichen

Merkmalen jedes Gliedes werden in erster Linie nur drei berücksichtigt, nämlich der beobachtete Zahlenwert, sodann der Stundenwinkel der Sonne oder die Tageszeit, endlich die Länge der Sonne in ihrer Bahn oder die Jahreszeit. Die beiden letztgenannten Merkmale dienen dazu, die Urliste auf verschiedene Arten in Kollektivreihen zu spalten, die als einziges Argument nur noch den beobachteten Zahlenwert enthalten. Anfangs beschränkte man sich nun darauf, nur das erste numerische Element jeder Kollektivreihe, nämlich das arithmetische Mittel der beobachteten Werte oder den Argumentdurchschnitt herzuleiten. Dann kam man zu der Einsicht, daß dieses Mittel als klimatologische Charakteristik zwar wesentlich und notwendig, aber für sich allein unzureichend sei; beispielsweise hat eine mittlere Januartemperatur von $+ 1.2^0$ eine sehr verschiedene Bedeutung, je nachdem die Einzelwerte stärker oder schwächer um den Mittelwert herum ausgestreut sind. Man ging deshalb dazu über, neben dem Argumentdurchschnitt auch Maßgrößen für die Streuung aufzustellen; die neben strengeren Methoden übliche Angabe der sogenannten Extreme und ihrer Mittelwerte ist in der Tat nichts anderes, als eine, allerdings rohe, Schätzung der Streuung.

Gegenwärtig ist man nun noch einen Schritt weiter gegangen. Konstruiert man nämlich aus den Kollektivreihen der Temperatur die Verteilungskurven, so haben letztere im allgemeinen die gewöhnliche Form, bei der ein Maximum mit Abfall nach beiden Seiten auftritt. Nun ist es aber für das Verständnis der meteorologischen Vorgänge durchaus nicht gleichgültig, ob die Kurven merklich symmetrisch oder merklich unsymmetrisch verlaufen, denn im letzteren Falle muß man schließen, daß die Ursachen, die eine Änderung erzeugen, zu beiden Seiten des Mittelwerts ungleich wirken. Deshalb werden jetzt auch die sogenannten „Gipfelwerte", d. h. die Maxima der Verteilungskurven, in die Untersuchung hineingezogen. Diese Berücksichtigung der Gipfelwerte ist aber nichts weiter, als eine weniger vollkommene Form für das, was bei unserer Darstellung der Kollektivreihen in präziser Gestalt durch die D-Koeffizienten zum Ausdruck gebracht wird.

Aus den vorstehenden Bemerkungen ersieht man, wie die in der Sache liegende Notwendigkeit von selber dahin drängt, daß die meteorologischen Kollektivreihen vollständiger und schärfer, als es seither geschah, den Grundsätzen der Kollektivmaßlehre unterworfen werden. Andererseits erkennt man aber auch, daß es für die Entwicklung der Klimatologie schwerlich von Vorteil gewesen wäre, wenn man von Anfang an auf die „Ausschöpfung" des Beobachtungsmaterials hätte ausgehen wollen. Das Gleiche gilt auch von anderen Gebieten, auf denen die Kollektivmaßlehre dazu berufen ist, als Untersuchungsmethode eine Rolle zu spielen: die Einsicht in das, was die numerischen Elemente niedrigerer Ordnung erkennen lassen,

ist zu Anfang wichtiger, als die Vollständigkeit des aufzustellenden Elementensystems.

§ 178. Die Φ-Reihe, von der wir seither ausgegangen sind, ist nur ein besonderer Fall aus einer unendlichen Mannigfaltigkeit von Darstellungsformen, die ebenfalls zur Grundlage der Kollektivmaßlehre gemacht werden könnten. Will man nun von dem besonderen Falle absehen und das schließliche Ziel der rechnerischen Behandlung allgemeiner formulieren, so läßt sich folgendes aussagen. Das zu erstrebende Ziel ist in der Kollektivmaßlehre wie in anderen messenden und zählenden Wissenschaften die *Verdichtung* jeder Beobachtungsreihe in eine Formel, die alles das ausspricht, was die Beobachtung an erkennbarer Gesetzmäßigkeit in sich schließt; man erreicht dadurch, daß die Beobachtung von der Formel vertreten werden kann. Sind also in unserem Falle zu einem vorgelegten K.-G. die beobachteten Werte der Summenfunktion $\mathfrak{S}(x)$ gebildet, so handelt es sich darum, einen analytischen Ausdruck $A(x)$ aufzustellen, der das beobachtete $\mathfrak{S}(x)$ nicht bloß im großen und ganzen, sondern, wie man kurz zu sagen pflegt, „erschöpfend“ wiedergibt. Hierzu ist keineswegs erforderlich, daß die zwischen Beobachtung und Rechnung übrigbleibenden Differenzen $\mathfrak{S} - A$ durchweg verschwindend klein ausfallen, denn die beobachteten Werte von $\mathfrak{S}(x)$ enthalten die Reste unausgeglichener Zufälligkeiten, deren Wiedergabe kein Interesse besitzt, weil ja $A(x)$ vor allem die in $\mathfrak{S}(x)$ vorhandenen Gesetzmäßigkeiten zum Ausdruck bringen soll; man wird im Gegenteil wünschen, daß jene Zufallsreste in $A(x)$ eine Ausgleichung erfahren. Darum genügt es für die sogenannte erschöpfende Darstellung, wenn die Differenzen $\mathfrak{S} - A$ nach Größe und Gruppierung so beschaffen sind, daß sie unbedenklich als die nicht weiter zu beseitigenden Überbleibsel zufälliger Umstände angesehen und deshalb außer Betracht gelassen werden können.

Sieht man sich nun die vorstehend formulierte Aufgabe genauer an, so ergibt sich, daß sie in drei Teilaufgaben zerfällt. Erstens nämlich ist die Gestalt der analytischen Funktion $A(x)$ festzustellen, die im allgemeinen gewisse, vorläufig unbekannte und aus den Beobachtungen zu ermittelnde Parameter enthalten wird. Zweitens sind, um den Anschluß von $A(x)$ an $\mathfrak{S}(x)$ zu bewerkstelligen, die numerischen Werte jener Parameter aufzusuchen. Drittens endlich hätte man Kriterien ausfindig zu machen, nach denen man zu beurteilen vermag, ob die übrigbleibenden Differenzen $\mathfrak{S} - A$ als die Reste unausgeglichener Zufälligkeiten angesehen werden dürfen oder nicht.

§ 179. Zu der ersten Teilaufgabe ist zu bemerken, daß es Fälle gibt, in denen die Gestalt der Funktion $A(x)$ von vornherein bestimmt und bekannt ist. Hat man z. B. Versuche mit Würfeln oder Glücksspielen oder anderen dem Urnenschema unterworfenen Vorgängen angestellt, so wird es gewöhnlich keine besonderen

Schwierigkeiten bieten, vermittelst der Lehrsätze der W.-R. die gesuchte Funktion $A(x)$ genau zu definieren und in geeigneter Weise analytisch auszudrücken. Diese Fälle bilden indessen nur eine bescheidene Minderheit: als Regel hat zu gelten, daß die Funktion $A(x)$ zunächst auch ihrer Form nach unbekannt ist. Da ferner $A(x)$ in den Fällen, wo man es sicher kennt, die allerverschiedensten Gestalten annehmen kann, so wird man bis zum Beweise des Gegenteils ruhig behaupten dürfen, daß das Gleiche auch in dem Gebiete der unbekannten $A(x)$ gelte. Aus diesem Grunde ist man — von besonderen Fällen abgesehen — in der Kollektivmaßlehre darauf angewiesen, bei der Ansetzung von $A(x)$ solche analytischen Gebilde zu benutzen, die eine *willkürlich* gegebene Summenfunktion darzustellen vermögen und zu dem Ende unendlich viele verfügbare Parameter enthalten. Das verlangte $A(x)$ wird also vermittelst eines unendlichen Prozesses aufgebaut, der selbstverständlich konvergent sein muß, da er andernfalls von vornherein unbrauchbar sein würde. Die Konvergenz bewirkt dann, daß bei den numerischen Anwendungen der Prozeß schon im endlichen abgebrochen werden darf.

Die Notwendigkeit, in der Kollektivmaßlehre von der Darstellung *willkürlicher* Verteilungen auszugehen, ist früher teils ganz übersehen, teils ungenügend berücksichtigt worden. Gleichwohl ist dieser Punkt wesentlich, denn wenn man für den darstellenden Ausdruck $A(x)$ von vornherein und allgemein eine *bestimmte* Funktion mit feststehender Parameterzahl zugrunde legt, so verfällt man in denselben prinzipiellen Fehler, den man früher beging, als man beobachtete Kollektivreihen in das gewöhnliche Exponentialgesetz hineinzuzwängen versuchte.

In dem bisherigen Gange unserer Untersuchung haben wir nun die Darstellung willkürlicher Verteilungen auf die Φ-Reihe gestützt, die aus der asymptotischen Gleichung

$$\operatorname{sg}(y-x) = \sum_n t^n \mathfrak{R}(x)_n \Phi(y)_n, \qquad (n = 0, 1, 2, \cdots)$$

entspringt. Diese Φ-Reihe werden wir auch jetzt ausschließlich benutzen, um die verlangten Funktionen $A(x)$ zusammenzusetzen. Damit ist selbstverständlich nicht ausgeschlossen, daß noch andere praktisch brauchbare und unter Umständen sogar vorteilhaftere Reihenentwicklungen vorhanden sein können; ebenso ist es sehr wohl möglich, daß die *Erfahrung* bei weiterer Entwicklung der angewandten Kollektivmaßlehre geschlossene Ausdrücke mit mäßiger Parameteranzahl kennen lehrt, die zwar nicht allgemein, wohl aber für gewisse Klassen von Kollektivreihen eine erschöpfende Darstellung liefern.

§ 180. Ist die Form des darstellenden Ausdruckes $A(x)$ festgesetzt, so hat als nächster Schritt die numerische Ermittelung der in $A(x)$ vorhandenen Parameter zu folgen. In unserem Falle treten als unbekannte Parameter die numerischen Elemente der Φ-Reihe

auf, wobei zu bemerken ist, daß die Berechnung dieser Elemente auch dann noch eine Bedeutung besitzt, wenn für $A(x)$ statt der Φ-Reihe ein geschlossener Ausdruck gegeben ist. Denn ein solches $A(x)$ läßt sich stets auch durch eine Φ-Reihe darstellen, nur daß dann die numerischen Elemente der Reihe bestimmte und bekannte Funktionen der in $A(x)$ enthaltenen Parameter sind. Man kann also aus den Elementen der Reihe die Parameter von $A(x)$ finden, wobei es für den Augenblick gleichgültig bleibt, ob ein derartiger Rechnungsgang vorteilhaft ist oder nicht.

In den früheren Abschnitten waren nun zwei verschiedene Wege zur Sprache gekommen, die bei der Berechnung der numerischen Elemente benutzt werden können, nämlich erstens die direkte Mittelbildung und zweitens die Auflösung eines Systems von linearen Gleichungen. Wir beschäftigen uns zunächst mit dem ersten Verfahren und wollen dabei der Einfachheit halber voraussetzen, daß es sich um einen unstetigen K-G. handele. Hierin liegt, wie wir bei der Untersuchung über den Einfluß der Abrundung gesehen haben, keine wesentliche Einschränkung der Allgemeinheit, da im Fall der Stetigkeit nur eine gewisse Reduktion erforderlich ist, die bei nicht zu starker Abrundung mit voller Sicherheit ermittelt werden kann.

In Übereinstimmung mit der seither benutzten Bezeichnungsweise soll x das Argument, $\mathfrak{U}(x)$ die Verteilungsfunktion, $\mathfrak{S}(x)$ die Summenfunktion und m den Umfang des vorgelegten K.-G. bedeuten. Die Verteilungstafel wird dann aus den zusammengehörigen Wertepaaren der Größen x und $\mathfrak{U}(x)$ gebildet, wobei die $\mathfrak{U}(x)$ aus den unmittelbar für das Argument x beobachteten Mengen $m\mathfrak{U}(x)$ durch Division mit m entstanden sind. Ob man bei der Aufsuchung der Elemente mit den ursprünglichen Größen $m\mathfrak{U}$ oder mit den daraus abgeleiteten $\mathfrak{U}$ rechnet, ist an sich gleichgültig, jedoch hat es in der Regel manches für sich, die beobachteten Zahlen $m\mathfrak{U}$ möglichst lange beizubehalten. Jedenfalls ist, wenn man die $\mathfrak{U}$ benutzt, darauf zu achten, daß ihre Summe streng den Wert Eins gibt, denn eine Abweichung hiervon, die ja durch Abrundung entstehen kann, ist stets lästig. Der zweckmäßigste Weg, die Summe der $\mathfrak{U}$ genau auf den Betrag Eins zu bringen, ist folgender. Man ergänzt das System der Halbierungspunkte, die zwischen den beobachteten x liegen, durch Hinzufügung der beiden Stellen $\pm\infty$, bildet für dieses vervollständigte System durch Aufsummieren der $m\mathfrak{U}(x)$ die Beträge der $m\mathfrak{S}(x)$ einschließlich der beiden Werte

$$m\mathfrak{S}(-\infty) = 0, \quad m\mathfrak{S}(+\infty) = m,$$

leitet daraus durch Division mit m die $\mathfrak{S}(x)$ selber her und gewinnt schließlich durch Subtraktion der aufeinanderfolgenden $\mathfrak{S}(x)$ die $\mathfrak{U}(x)$. Bei diesem Verfahren wird offenbar die Summe der $\mathfrak{U}$ streng gleich Eins. In ähnlicher Weise wird man, wenn etwa die beobachteten

$m\mathfrak{U}$ mit den aus der Φ-Reihe berechneten Zahlen verglichen werden sollen, die theoretischen $m\mathfrak{U}$ nicht durch Multiplikation der berechneten $\mathfrak{U}$ mit m bilden, sondern durch Subtraktion aus den vermittelst der Φ-Reihe berechneten Werten von $m\mathfrak{S}(x)$.

§ 181. Bei der Berechnung der Elemente kann man sich zunächst an die Gleichungen halten, durch die die einzelnen Elemente definiert werden. Danach hat man, wenn mit den beobachteten Mengen $m\mathfrak{U}$ gerechnet wird, mit der Bildung der über alle beobachteten x erstreckten Produktsumme

$$\sum_x x \cdot m\mathfrak{U}(x) = m\mathfrak{D}(x)$$

zu beginnen, woraus das erste Element $c = \mathfrak{D}(x)$ folgt. Dann setzt man die Größen $y = x - c$ nebst deren Potenzen $y^2, y^3, \ldots$ an und und bildet die Produktsummen

$$\sum_x y^n m\mathfrak{U}(x) = m\mathfrak{D}(y^n), \tag{1}$$

in denen n der Reihe nach gleich $2, 3, \ldots p$ zu nehmen ist, wenn D_p den höchsten noch mitzunehmenden D-Koeffizienten bedeutet. Die Summe für $n = 2$ liefert den Wert von $m \operatorname{str}(x)^2$, woraus das zweite Element $\operatorname{str}(x)$ und mit

$$2h^2 \operatorname{str}(x)^2 = 1$$

der normale Wert des Parameters h folgt. Multipliziert man ferner die für $n = 3, 4, \ldots$ gefundenen Produktsummen mit $h^3, h^4, \ldots$ so ergeben sich die Werte von

$$m\mathfrak{D}[(hy)^3], \quad m\mathfrak{D}[(hy)^4], \quad \ldots,$$

aus denen nach Division mit m die noch fehlenden Elemente $D_3, D_4, \ldots$ als lineare Verbindungen mit festen Zahlenkoeffizienten zusammenzusetzen sind.

Ein zweiter Weg ergibt sich, wenn man nicht sogleich auf die Normalform ausgeht. Mit der irgendwie gewählten Größe e bilde man die Werte von $z = x - e$ und dazu die Potenzen von z. Dann berechne man die über alle beobachteten x erstreckten Produktsummen

$$\sum_x z^n m\mathfrak{U}(x) = m\mathfrak{D}(z^n), \tag{2}$$

für $n = 1, 2, 3, \cdots$ und bilde aus den beiden ersten Summen die Elemente $\mathfrak{D}(x)$ und $\operatorname{str}(x)$ nach den Gleichungen

$$\mathfrak{D}(z) = \mathfrak{D}(x) - e, \quad \operatorname{str}(x)^2 = \mathfrak{D}(z^2) - \mathfrak{D}(z)^2.$$

Ferner rechnet man die Größen $h^n\mathfrak{D}(z^n)$ mit dem normalen h und findet daraus die Durchschnitte

$$\mathfrak{D}[\mathfrak{R}(hx - he)_n], \tag{3}$$

die schließlich noch nach den Formeln in § 99 auf die normalen Werte

$$D_n = \mathfrak{D}[\mathfrak{R}(hx - hc)_n]$$

zu reduzieren sind.

Das vorstehende zweite Verfahren erscheint zunächst als ein völlig überflüssiger Umweg, da man die anfangs eingeführte Größe e nachträglich bei dem Übergange auf die Normalform wieder fortzuschaffen hat. Die Sache gewinnt jedoch sofort ein anderes Gesicht, wenn die x ganzzahlig sind und eine größere Anzahl von Kollektivreihen zu bearbeiten ist. Denn dann werden, wenn man für e eine ganze Zahl wählt, die z ebenfalls ganze Zahlen; man braucht also die Tabelle der z^n nur einmal zu berechnen und kann sie nötigenfalls in einer für die Rechnung bequemen Gestalt vervielfältigen lassen. Dadurch wird die Mehrarbeit, die aus dem Übergange von (3) auf die Normalform erwächst, reichlich wieder eingebracht.

Rechnet man auf einem der beiden hier skizzierten Wege ein längeres Beispiel durch, so erkennt man bald, daß der lästigste Teil der Arbeit die Bildung der Produkte $y^n m\mathfrak{U}$ oder $z^n m\mathfrak{U}$ ist. Die Ursache davon liegt weniger in der Menge der auszuführenden Multiplikationen, als vielmehr darin, daß die einzelnen Produkte in der Hauptsache unabhängig voneinander entstehen. Infolgedessen ist man behufs durchgreifender Kontrolle der Zahlen zu einer Wiederholung der Rechnung genötigt; wenigstens ist es mir nicht gelungen, für die beiden angeführten Methoden Kontrollgrößen ausfindig zu machen, deren Benutzung vor der doppelten Rechnung allgemein den Vorzug verdiente. Unter diesen Umständen ist es von Belang, daß man noch ein drittes Verfahren entwickeln kann, das zwar mit der Unbequemlichkeit eines stärkeren Ziffernverbrauchs behaftet ist, dafür aber andere wertvolle Vorzüge besitzt, die jenen Mangel mehr als aufwiegen. Dieses Verfahren, das man kurz als die *Summenmethode* bezeichnen kann, soll nunmehr behandelt werden.

Einundzwanzigste Vorlesung.

Numerische Bearbeitung: Summenmethode (I. Form).

§ 182. Die Summenmethode, mit der wir uns jetzt zu beschäftigen haben, beruht auf dem Verhalten der sogenannten Summenreihen. Diese Reihen haben auf einem andern Gebiete des wissenschaftlichen Rechnens, nämlich bei der numerischen Integration, schon seit bald einem Jahrhundert eine ausgiebige Verwendung gefunden, die allerdings außerhalb des Kreises der berufsmäßigen Rechner von jeher

ziemlich unbekannt geblieben ist. Charakteristisch für die hierher gehörigen Methoden ist, daß sie vorwiegend nicht mit geschlossenen fertigen Ausdrücken arbeiten, sondern mit einfachen und leicht zu beschreibenden Rechnungsprozessen, wie das im einzelnen weiterhin zum Vorschein kommen wird.

Der Gebrauch der Summenmethode ist an eine Bedingung gebunden, die bei beobachteten Kollektivreihen nur ausnahmsweise nicht erfüllt ist und die wir mit Rücksicht auf den weiterhin befolgten Rechnungsgang als „Einführung des Nummernarguments" bezeichnen können. Es ist zunächst anzugeben, um was es sich hierbei handelt.

Die weitaus überwiegende Mehrheit der beobachteten Verteilungstafeln ist so beschaffen, daß sich die Abstände zwischen den notierten Argumentwerten als ganzzahlige Vielfache einer Größe b darstellen lassen, die z. B. bei den stetigen Reihen mit äquidistanten Wechselpunkten durch die konstante Teilstreckenlänge gegeben ist. Bezeichnet dann a irgend ein notiertes Argument, so ist für jedes andere notierte Argument x die Differenz $x - a$ ein Vielfaches von b, und man darf demgemäß ansetzen

$$x = a + bX, \tag{1}$$

wobei X immer eine ganze Zahl bedeutet. Verschiebt man den Punkt a um ein Vielfaches von b, so ändern sich die X gleichmäßig um eine bestimmte ganze Zahl. Man kann also a so wählen, daß zu dem niedrigsten vollen x der Wert $X = 1$, und zu den nachfolgenden x die Werte aus der Reihe $X = 2, 3, \ldots$ gehören. Ebenso kann man, da negative Werte von b nicht ausgeschlossen sind, a auch so wählen, daß zu dem größten vollen x wiederum der Wert $X = 1$, und zu den vorangehenden x die Werte aus der Reihe $X = 2, 3, \ldots$ gehören. Das so bestimmte neue Argument X wollen wir fortan als *Nummernargument* bezeichnen. Ferner setzen wir voraus, daß das Argument des zu untersuchenden K.-G. — wenn nötig durch eine vorhergegangene Transformation nach (1) — in die Form des Nummernarguments gebracht worden sei, daß also die Kollektivreihen, bei denen eine solche Transformation nicht angängig ist, von der Untersuchung ausgeschlossen bleiben.

§ 183. Bei der Herleitung der Summenmethode gehen wir von dem nachstehenden und sogleich zu erläuternden Größenschema aus:

$$\begin{array}{ccccccc}
\ldots & a(-2) & a(-1) & a(0) & a(1) & a(2) & \ldots \\
\ldots & b(-2) & b(-1) & b(0) & b(1) & b(2) & \ldots \\
\ldots & c(-2) & c(-1) & c(0) & c(1) & c(2) & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots
\end{array}$$

Die Zeichen $a(h)$, deren Reihe nach rechts und links vorläufig nicht begrenzt sein soll, bedeuten gegebene Zahlen, die weiterhin immer

ganz und positiv sein werden. Die Größe $b(h)$ entsteht, wenn man alle links von $a(h)$ stehenden a-Werte summiert, d. h. es ist

$$b(h) = a(h-1) + a(h-2) + a(h-3) + \cdots. \qquad (2)$$

Ebenso entsteht $c(h)$ durch Summation der links von $b(h)$ stehenden b-Werte, so daß

$$c(h) = b(h-1) + b(h-2) + b(h-3) + \cdots \qquad (3)$$

wird. In dieser Weise geht es weiter zu den folgenden Zeilen des Schemas.

Die $b(h)$, $c(h)$, ... liefern die sogenanten *Summengrößen erster, zweiter usw. Ordnung.*

Aus den Definitionsgleichungen der $b(h)$, $c(h)$, ... folgt sofort

$$b(h+1) - b(h) = a(h),$$
$$c(h+1) - c(h) = b(h),$$
usw.

Sind also z. B. außer den $a(h)$ die Werte von $b(1)$, $c(1)$, ... gegeben, so entstehen die darauf folgenden Werte durch die Operationen

$$b(2) = b(1) + a(1), \quad b(3) = b(2) + a(2), \quad \ldots \qquad (4)$$
$$c(2) = c(1) + b(1), \quad c(3) = c(2) + b(2), \quad \ldots \qquad (5)$$
usw.

Setzt man für die $a(h)$ durchweg positive ganze Zahlen (mit Einschluß der Null), so müssen, damit die $b(h)$, $c(h)$, ... nicht unendlich werden, die $a(h)$ links von einer gewissen Nummer beständig null sein. Wir wollen demgemäß festsetzen, daß die Größen

$$a(0), \quad a(-1), \quad a(-2), \quad \ldots$$

durchweg verschwinden. Es wird dann

$$b(1) = 0, \quad c(1) = 0, \quad \text{usw.},$$

so daß die vorhin angegebene Rechnung sofort beginnen kann.

Über die Fortsetzung des Schemas nach rechts ist zu bemerken, daß bei der numerischen Bearbeitung von Kollektivreihen die rechts von einer gewissen Nummer stehenden $a(h)$ beständig verschwinden, und daß ferner, wenn $a(k)$ das letzte von Null verschiedene $a(h)$ bedeutet, die Rechnung nur bis $b(k+1)$, $c(k+1)$, ... durchzuführen ist.

§ 184. Um den vorstehenden Ansatz noch deutlicher zu machen, möge als Beispiel die folgende, 14 Glieder umfassende Reihe der Größen $a(1)$, ... behandelt werden.

$$\begin{array}{ccccccc} 6 & 36 & 78 & 149 & 161 & 183 & 134 \\ 114 & 74 & 34 & 19 & 10 & 0 & 2 \end{array} \qquad (6)$$

Die nachstehende Tabelle I enthält den Anfang der Rechnung. Die Nullspalten mit $a(0)$, $a(-1)$ usw. sind, da sie jetzt keinen Zweck mehr haben, einfach fortgelassen. Die Zeile über dem Strich gibt die Spaltennummern an, ferner sind linker Hand die Zeilenbuchstaben $a, b, \ldots$ angesetzt.

Tabelle I.

	1	2	3	4	5	...
a	6	36	78	149	161	...
b	0	6	42	120	269	...
c	0	0	6	48	168	...
d	0	0	0	6	54	...

Nachdem man das Rechenschema durch Eintragung der Spaltennummern und der Zeilenbuchstaben vorbereitet hat, trägt man in die a-Zeile die a-Größen aus (6) ein. Darauf setzt man in der b-Zeile den Anfangswert $b(1) = 0$ an und rechnet nun nach (4)

$$b(2) = b(1) + a(1) = 0 + 6 = 6,$$
$$b(3) = b(2) + a(2) = 6 + 36 = 42,$$
usw.

Ist die b-Zeile fertig, so trägt man auf der c-Zeile das Anfangsglied $c(1) = 0$ ein und rechnet nach (5)

$$c(2) = c(1) + b(1) = 0 + 0 = 0,$$
$$c(3) = c(2) + b(2) = 0 + 6 = 6,$$
usw.

In derselben Weise entsteht dann die d-Reihe aus dem Anfangsgliede $d(1) = 0$.

Die beschriebene Anordnung der Rechnung ist an sich hinreichend einfach und übersichtlich, aber nicht sonderlich bequem für die bei solchen Arbeiten unentbehrliche Kontrolle. Aus diesem Grunde empfiehlt es sich, die in Tabelle I benutzten vierzeiligen Spalten nicht neben, sondern untereinander zu setzen, soweit es das Papierformat zuläßt. Eine Vorstellung hiervon gibt die Tabelle II, in welcher jedoch des Druckes wegen, 3 Spalten zu 20 Zeilen statt einer einzigen Spalte zu 60 Zeilen angesetzt worden sind.

Das Schema enthält 15 Felder, von denen die 14 ersten den Spalten entsprechen, die in Tabelle I aus den 14 in (6) enthaltenen a-Größen entstehen. Das 15-te Feld dient dann für die Größen $b(15)$, $c(15)$, $d(15)$, bis zu denen im vorliegenden Falle die Rechnung gehen soll. Jedes Feld enthält linker Hand die Feldnummer, die mit der Spaltennummer in Tabelle I übereinstimmt, und daneben die Zeilenbuchstaben.

Tabelle II.

1.a	6	6.a	183	11.a	19
1.b	0	6.b	430	11.b	969
1.c	0	6.c	437	11.c	4 023
1.d	0	6.d	222	11.d	8 321
2.a	36	7.a	134	12.a	10
2.b	6	7.b	613	12.b	988
2.c	0	7.c	867	12.c	4 992
2.d	0	7.d	659	12.d	12 344
3.a	78	8.a	114	13.a	0
3.b	42	8.b	747	13.b	998
3.c	6	8.c	1480	13.c	5 980
3.d	0	8.d	1526	13.d	17 336
4.a	149	9.a	74	14.a	2
4.b	120	9.b	861	14.b	998
4.c	48	9.c	2227	14.c	6 978
4.d	6	9.d	3006	14.d	23 316
5.a	161	10.a	34	15.a	—
5.b	269	10.b	935	15.b	1 000
5.c	168	10.c	3088	15.c	7.976
5.d	54	10.d	5233	15.d	30 294

Nachdem das Schema mit der Feldereinteilung vorbereitet ist, trägt man auf den a-Zeilen die in (6) angesetzten Größen ein und sperrt der Sicherheit halber die Zeile 15.a durch einen Strich Darauf summiert man in *einem Zuge* die eingetragenen $a(h)$, setzt die Summe in 15.b und kontrolliert die Summation sogleich durch Wiederholung in umgekehrter Richtung. Nach dieser Vorbereitung wird in 1.b der Anfangswert $b(1) = 0$ eingetragen und schrittweise nach der Vorschrift

$$2.b = 1.b + 1.a, \qquad 3.b = 2.b + 2.a, \quad \text{usw.}$$

gerechnet. Das Ende der Rechnung muß auf die bereits in $15.b$ stehende Endsumme führen, womit offenbar alle $b(h)$ kontrolliert sind. Jetzt werden weiter die $b(h)$ von $h = 1$ bis $h = 14$ in einem Zuge summiert: die Summe kommt nach $15.c$ und wird sofort nachgeprüft. Darauf trägt man in $1 \cdot c$ den Anfangswert $c(1) = 0$ ein und berechnet schrittweise die folgenden c-Größen, die wieder durch die bereits vorhandene Endsumme kontrolliert werden. In dieser Weise geht es weiter, bis alle Summenreihen, die man braucht, gefunden sind.

Die Anzahl der Summenreihen, die man zu berechnen hat, hängt von der Zahl der Glieder ab, die in der Φ-Reihe berücksichtigt werden

sollen. Es wird sich zeigen, daß man bis zu den Summen der Ordnung $p+1$ zu gehen hat, wenn man die D-Koeffizienten bis $D(c, h)_p$ ermitteln will. Dieser Umstand ist bei der Anlegung des Felderschemas zu berücksichtigen. Ferner kommt man, wenn die $a(h)$ ganze positive Zahlen sind, bei den höheren Summengrößen auf vielstellige Zahlen, was man ebenfalls bei der Anlegung des Schemas zu beachten hat. Hierbei mag als eine unscheinbare, aber praktisch keinesweg unwichtige Äußerlichkeit erwähnt werden, daß es zweckmäßig ist, die Zahlen etwas weitläufig zu schreiben, da sich andernfalls bei den in einem Zuge auszuführenden Summationen das Auge leicht „verlaufen" kann.

§ 185. Nachdem wir das Rechenschema für die Summengrößen kennen gelernt haben, ist nunmehr der Zusammenhang zwischen diesen Größen und den gegebenen Zahlen $a(h)$ zu ermitteln. Zu dem Ende soll zuerst die Bezeichnungsweise etwas abgeändert werden.

Der auf ganzzahlige Werte beschränkten Veränderlichen X werde die davon abhängende Größe $a(X)$ zugeordnet. Wenn X die Zahlen $1, 2, \ldots n$ durchläuft, so soll $a(X)$ vorgeschriebene, im allgemeinen von Null verschiedene Werte annehmen. Dagegen soll $a(X)$ für die Argumentreihen

$$X = 0, -1, -2, \ldots$$
$$X = n+1, n+2, n+3, \ldots$$

beständig verschwinden. Dies festgesetzt bilde man die Summengrößen

$$(X, 1) = a(X-1) + a(X-2) + \cdots,$$
$$(X, 2) = (X-1, 1) + (X-2, 1) + \cdots,$$
$$(X, 3) = (X-1, 2) + (X-2, 2) + \cdots,$$
usw.

Hierin entsprechen die $a(X)$ den vorhin mit $a(h)$ bezeichneten Zahlen, während die (X, p) die aus den $a(X)$ entspringenden Summengrößen p-ter Ordnung bedeuten. Aus den Definitionsgleichungen folgt

$$(X+1, p+1) - (X, p+1) = (X, p). \tag{7}$$

Hieraus würde für $p = 0$ die Beziehung

$$(X+1, 1) - (X, 1) = (X, 0)$$

fließen, der zufolge das vorläufig nicht definierte Zeichen $(X, 0)$ als mit $a(X)$ zusammenfallend anzunehmen ist; man kann also die $a(X)$ als die Summengrößen nullter Ordnung auffassen.

Weiter werde mit der Veränderlichen t die Reihe

$$L(X) = (X, 1) + (X, 2)t + (X, 3)t^2 + \cdots$$

oder kürzer

$$L(X) = \sum_p (X, p+1)t^p, \qquad (p = 0, 1, \ldots) \tag{8}$$

gebildet. Multipliziert man darauf (7) mit t^p und summiert das Produkt nach p von $p = 0$ an, so entsteht

$$L(X+1) - L(X) = a(X) + tL(X)$$

oder

$$L(X+1) - (1+t)L(X) = a(X)$$

und daraus

$$(1+t)^{-X-1}L(X+1) - (1+t)^{-X}L(X) = a(X)(1+t)^{-X-1},$$

Ersetzt man hierin X durch $1, 2, \ldots Y$ und summiert die entstehenden Gleichungen, so wird

$$(1+t)^{-Y-1}L(Y+1) - (1+t)^{-1}L(1) = \sum_X a(X)(1+t)^{-X-1}, \quad (X = 1, 2, \ldots Y).$$

Nach den Definitionsgleichungen der (X, p) werden aber wegen des Verhaltens der $a(X)$ die Größen $(1, 1)$, $(1, 2)$, ... beständig null, und das Gleiche gilt wegen (8) von $L(1)$, so daß man

$$L(Y+1) = \sum_X a(X)(1+t)^{Y-X}$$

erhält. Ersetzt man den Summationsbuchstaben X durch $Y-Z$, so wird

$$L(Y+1) = \sum_Z a(Y-Z)(1+t)^Z, \tag{9}$$

wobei Z von 0 bis $Y-1$ zu laufen hat, jedoch noch darüber hinaus fortgesetzt werden kann, da $a(Y-Z)$ für $Z = Y, Y+1, \ldots$ durchweg verschwindet. Führt man nun die Binomialreihe

$$(1+t)^Z = (Z)_0 + (Z)_1 t + (Z)_2 t^2 + \cdots = \sum_p (Z)_p t^p$$

ein und ersetzt $L(Y+1)$ durch die zugehörige Reihe, so wird

$$\sum_p (Y+1, p+1)t^p = \sum_Z \sum_p a(Y-Z)(Z)_p t^p,$$

woraus durch Spaltung nach den Potenzen von t die Relation

$$(Y+1, p+1) = \sum_Z a(Y-Z)(Z)_p \tag{10}$$

folgt. Damit sind die Summengrößen explizite durch die $a(X)$ ausgedrückt. Aus den Summen in (10) sind nun zunächst die Summen über die Produkte $a(Y-Z)Z^p$ herzuleiten. Hierzu dient ein Prozeß, den man füglich als das Rechnen mit *Paarsummen* bezeichnen kann.

§ 186. Für die Binomialkoeffizienten gilt der Ausdruck

$$(Z)_p = Z! : [p!\,(Z-p)!],$$

aus dem

$$(Z)_{p+1} : (Z)_p = (Z-p) : (p+1)$$

oder

$$(p+1)(Z)_{p+1} + p(Z)_p = Z(Z)_p \tag{11}$$

folgt. Man denke sich nun die Reihe der Binomialkoeffizienten

$$(Z)_0, \quad (Z)_1, \quad (Z)_2, \quad (Z)_3, \quad \ldots \tag{12}$$

hingeschrieben und darauf die Glieder dieser Reihe mit 0, 1, 2, 3, ... oder ihrem Index multipliziert, so daß die Reihe der Produkte

$$0 \cdot (Z)_0, \quad 1 \cdot (Z)_1, \quad 2 \cdot (Z)_2, \quad 3 \cdot (Z)_3, \ldots \tag{13}$$

entsteht. Addiert man jetzt je zwei benachbarte Produkte, so nehmen diese *Paarsummen* wegen (11) die Werte

$$Z(Z)_0, \quad Z(Z)_1, \quad Z(Z)_2, \ldots \tag{14}$$

an, d. h. es entsteht wieder die Reihe (12), jedoch multipliziert mit dem dem Faktor Z. Wiederholt man nun den Prozeß an (14), so entsteht die Reihe

$$Z^2(Z)_0, \quad Z^2(Z)_1, \quad Z^2(Z)_2, \ldots,$$

und so geht es beliebig weiter.

Der beschriebene Prozeß soll nun auch noch an den durch (10) bestimmten Größen $(Y+1, p+1)$ für $p = 0, 1, \ldots$ ausgeführt werden. Man hat also anzusetzen

$$(Y+1, 1), \quad (Y+1, 2), \quad (Y+1, 3), \ \ldots$$

dann die Produkte

$$0 \cdot (Y+1, 1), \quad 1 \cdot (Y+1, 2), \quad 2 \cdot (Y+1, 3), \ldots$$

und daraus die Paarsummen zu bilden, die mit

$$(Y+1, 1)_1, \quad (Y+1, 2)_1, \quad (Y+1, 3)_1, \ldots$$

bezeichnet werden sollen. Der zweite Schritt liefert die Produkte

$$0 \cdot (Y+1, 1)_1, \quad 1 \cdot (Y+1, 2)_1, \quad 2 \cdot (Y+1, 3)_1, \ldots$$

und dazu die Paarsummen

$$(Y+1, 1)_2, \quad (Y+1, 2)_2, \quad (Y+1, 3)_2, \ldots,$$

die dann in derselben Weise weiter zu bearbeiten sind. Dieser Prozeß bricht von selber ab, wenn — wie es bei den Anwendungen der Fall ist — die Summengrößen nur bis zu einer bestimmten Ordnungsnummer berechnet vorliegen.

Führt man endlich noch den Prozeß gleichzeitig q-mal an den beiden Seiten der Gleichungen aus, die aus (10) durch die Substitution $p = 0, 1, \ldots$ entstehen, so erhält man die allgemeine Beziehung

$$(Y+1, p+1)_q = \sum_Z a(Y-Z) Z^q (Z)_p. \tag{15}$$

Diese Gleichung gilt auch für $q = 0$, wenn man unter $(Y+1, p+1)_0$ die Ausgangsgrößen $(Y+1, p+1)$ versteht und auf der rechten Seite Z^0 durch den Wert 1 auch in dem Falle $Z = 0$ ersetzt. Für $p = 0$ wird

$$(Y+1, 1)_q = \sum_Z a(Y-Z) Z^q, \tag{16}$$

d. h. die Anfangsglieder der aus den $(Y+1, p+1)$ gebildeten Paar-

summenreihen liefern die Werte der auf der rechten Seite von (16) stehenden Potenzsummen.

§ 187. Die bisher entwickelten Formeln sollen jetzt zur Herleitung derjenigen Rechenvorschriften dienen, die ich als die *erste Form der Summenmethode* bezeichnen will. Zu dem Ende denken wir uns einen unstetigen K.-G. mit dem Umfange m und der Verteilungsfunktion $\mathfrak{U}(x)$ gegeben, dessen Argument — nötigenfalls nach einer vorhergegangenen Transformation — nur ganzzahlige Werte annimmt. Hierbei sollen g und k das größte und das kleinste volle Argument bedeuten und

$$n = g - k + 1 \tag{17}$$

gesetzt werden. Daneben führen wir das Nummernargument X durch die Gleichung

$$X = x - k + 1 \tag{18}$$

ein, so daß X von 1 bis n zu laufen hat, wenn man die vollständige Reihe der für $\mathfrak{U}(x)$ in Betracht kommenden x von $x = k$ bis $x = g$ erhalten will. Setzt man nun für die vorhin eingeführten Größen $a(1), \ldots a(n)$ der Reihe nach die Werte von $m\mathfrak{U}(x)$, die ihrer Natur nach ganzzahlig sind, so ist

$$a(1) = m\mathfrak{U}(k), \quad a(2) = m\mathfrak{U}(k+1), \quad \ldots \quad a(n) = m\mathfrak{U}(g) \tag{19}$$

oder zusammenfassend

$$a(X) = a(x - k + 1) = m\mathfrak{U}(x). \tag{20}$$

Da die $\mathfrak{U}(x)$ für die x unterhalb k und oberhalb g verschwinden, so gilt gleiches von den $a(X)$ für die X unterhalb 1 und oberhalb n, wie es sein soll. Damit liefert die Gleichung (16), wenn $Y = n$ gesetzt wird, die Beziehung

$$(n+1,\ 1)_q = \sum_Z a(n - Z) Z^q$$

oder mit $Z = g - x$

$$(n+1,\ 1)_q = \sum_x a(x - k + 1)(g - x)^q,$$

woraus wegen (20)

$$(n+1,\ 1)_q = m \sum_x \mathfrak{U}(x)(g - x)^q$$

oder

$$m\mathfrak{D}[(g - x)^q] = (n+1,\ 1)_q \tag{21}$$

folgt. Berechnet man also die Summengrößen mit den in (19) angesetzten n Größen $a(X)$ bis zur Spalten- oder Feldnummer $n+1$, so ergeben die daraus durch den Paarsummenprozeß abgeleiteten Größen $(n+1,\ 1)_q$ die in (21) angesetzten D-Größen, aus denen dann noch die $\mathfrak{D}(x^q)$ und die numerischen Elemente herzuleiten sind.

Statt des Ansatzes (18) und (20) kann man für das Nummernargument auch die Gleichungen

$$\begin{aligned} X &= g - x + 1, \\ a(X) &= a(g - x + 1) = m\mathfrak{U}(x) \end{aligned} \tag{22}$$

zugrunde legen, nach denen die $a(X)$ mit den in umgehrter Reihenfolge genommenen $m\mathfrak{U}(x)$ übereinstimmen, da ja nach (22)

$$a(1) = m\mathfrak{U}(g), \quad a(2) = m\mathfrak{U}(g-1), \cdots a(n) = m\mathfrak{U}(k) \tag{23}$$

wird. Man erhält dann aus (16) für $Y = n$ zunächst

$$(n+1,\ 1)_q = \sum_Z a(n-Z) Z^q,$$

daraus mit $Z = x - k$

$$(n+1,\ 1)_q = \sum_x a(g-x+1)(x-k)^q = m \sum_x \mathfrak{U}(x)(x-k)^q$$

oder

$$m\mathfrak{D}[(x-k)^q] = (n+1,\ 1)_q. \tag{24}$$

Aus (21) und (24) folgt noch für $q = 0$, daß die erste Summengröße, nämlich $(n+1,\ 1)$, jedesmal den Umfang der Kollektivreihe liefert.

§ 188. Bevor die Entwicklung der Rechenvorschriften weiter fortgesetzt wird, möge das bisherige Ergebnis an einem numerischen Beispiele erläutert werden.

Der „Thesaurus logarithmorum etc.“ von *Vega* enthält in seinem ersten Teile die zehnstelligen Logarithmen für die Numeri von 10000 bis 101000. Die Logarithmen sind auf jeder Seite des Werkes in 5 Spalten zu je 60 Zeilen untergebracht. Schon vor langen Jahren hatte ich nun die ersten 1000 Spalten daraufhin ausgezählt, wie oft sie eine *Endnull*, d. h. eine Null in der letzten Stelle enthalten. Die Zählung lieferte offenbar einen K.-G. vom Umfange $m = 1000$, dessen Glieder die genannten Spalten, und dessen Argument die Menge der Endnullen jeder Spalte sind. Das Ergebnis war die nachstehende Verteilungstafel, die keiner besondern Erläuterung bedarf.

Tabelle III.

$x =$	1	2	3	4	5	6	7
$m\mathfrak{U}(x) =$	6	36	78	149	161	183	134
$x =$	8	9	10	11	12	13	14
$m\mathfrak{U}(x) =$	114	74	34	19	10	0	2

Hiernach kommt der Fall $x = 0$ nicht vor, ebenso fehlen die x oberhalb 14. Eine Transformation auf ganzzahliges Argument ist nicht erst nötig, und man hat

$$k = 1, \quad g = 14, \quad n = 14.$$

Die $m\mathfrak{U}(x)$ sind nun identisch mit den oben in (6) angeführten Zahlen, die in Tabelle II weiter bearbeitet wurden. Demnach können wir aus Tabelle II unmittelbar die für den Ansatz (20) geltenden Summengrößen ausschreiben und erhalten

$$(n+1, 1) = 1000, \quad (n+1, 2) = 7976, \quad (n+1, 3) = 30294.$$

Daraus entspringen die Produkte

$$0, \quad 7976, \quad 60588$$

nebst den Paarsummen

$$(n+1, 1)_1 = 7976, \quad (n+1, 2)_1 = 68564$$

und weiter

$$(n+1, 1)_2 = 68564.$$

Demnach ist nach (21)

$$m\mathfrak{D}[14-x] = 7976, \quad m\mathfrak{D}[(14-x)^2] = 68564. \tag{25}$$

Rechnet man nach dem Ansatz (22), so sind in Tabelle II die $m\mathfrak{U}(x)$ in umgekehrter Reihenfolge einzutragen. Mit Unterdrückung der Zwischenrechnung erhält man

$$(n+1, 1) = 1000, \quad (n+1, 2) = 5024, \quad (n+1, 3) = 12582,$$

daraus

$$(n+1, 1)_1 = 5024, \quad (n+1, 1)_2 = 30188$$

und weiter nach (24)

$$m\mathfrak{D}[x-1] = 5024, \quad m\mathfrak{D}[(x-1)^2] = 30188. \tag{26}$$

Da nun für $y = x - g$ oder $y = x - k$ die Streuung aus

$$\operatorname{str}(x)^2 = \mathfrak{D}(y^2) - \mathfrak{D}(y)^2$$

gefunden wird, so erhält man mit $m = 1000$ aus (25) und (26) übereinstimmend

$$\mathfrak{D}(x) = 6.024, \quad \operatorname{str}(x)^2 = 4.947, \quad \operatorname{str}(x) = 2.224. \tag{27}$$

Wenn die Zahlen n und m größere Werte besitzen, so treten unter den Ausdrücken (21) und (24) vielstellige Zahlen auf, die man vor der weiteren Benutzung zweckmäßigerweise erst reduziert. Hierzu dient ein Verfahren, das wiederum auf der Verwendung der Paarsummen beruht und jetzt dargelegt werden soll.

§ 189. Man denke sich für die Konstante c und für $y = x - c$ die Größen

$$m\mathfrak{D}(y^0), \quad m\mathfrak{D}(y^1), \quad m\mathfrak{D}(y^2), \ldots \tag{28}$$

bis zu einem gewissen Exponenten hin gegeben, dann kann man sich die Aufgabe stellen, aus (28) auf einem möglichst einfachen Wege die mit der Konstanten d und mit $z = x - d$ gebildeten Größen

$$m\mathfrak{D}(z^0), \quad m\mathfrak{D}(z^1), \quad m\mathfrak{D}(z^2), \ldots \tag{29}$$

zu berechnen. Setzt man nun $e = c - d$ und addiert zu jedem Gliede in (28) das mit e multiplizierte vorhergehende Glied, so liefern diese Paarsummen, wie leicht zu sehen, die Werte von

$$m\mathfrak{D}(z), \quad m\mathfrak{D}(zy), \quad m\mathfrak{D}(zy^2), \ldots.$$

Wird an den gefundenen Zahlen der beschriebene Prozeß wiederholt, so entstehen die Werte von

$$m\mathfrak{D}(z^2), \quad m\mathfrak{D}(z^2y), \quad m\mathfrak{D}(z^2y^2), \ldots.$$

Wird der Prozeß fortgesetzt, bis er von selber abbricht, so sind die Anfangsglieder der einzelnen Paarsummenreihen nichts anderes, als die gesuchten Größen (29). Hiernach würde also für den betrachteten K.-G. und für $c = 14$, $d = 6$, $e = 8$ nach den Zahlen (25) die Rechnung folgendermaßen verlaufen:

$$\begin{array}{rrrr}
q = & 0 & 1 & 2 \\
\hline
m\mathfrak{D}(y^q) = & +1000 & -7976 & +68554 \\
 & & +8000 & -63808 \\
\cline{3-4}
 & & +\quad 24 & +\ 4756 \\
 & & & +\quad 192 \\
\cline{4-4}
 & & & +\ 4948
\end{array}$$

woraus zu $z = x - 6$ die Werte

$$\mathfrak{D}(z) = 0.024, \qquad \mathfrak{D}(z^2) = 4.948, \qquad \mathrm{str}\,(x)^2 = 4.947 \tag{30}$$

folgen. In derselben Weise wird mit den Zahlen (26) und $c = 1$, $d = 6$, $e = -5$

$$\begin{array}{rrrr}
q = & 0 & 1 & 2 \\
\hline
m\mathfrak{D}(y^q) = & +1000 & +5024 & +30188 \\
 & & -5000 & -25120 \\
\cline{3-4}
 & & +\quad 24 & +\ 5068 \\
 & & & -\quad 120 \\
\cline{4-4}
 & & & +\ 4948
\end{array}$$

womit man wieder auf die Werte in (30) kommt.

Die beschriebene Methode läßt, wenn es nur auf die Elemente $\mathfrak{D}(x)$ und $\mathrm{str}\,(x)$ ankommt, an Bequemlichkeit nichts zu wünschen übrig und führt auch bei größeren n rasch genug zum Ziele, wobei ein Hauptvorzug vor der früher besprochenen direkten Mittelbildung in den sich von selbst darbietenden Kontrollen liegt. Sollen dagegen die numerischen Elemente über $\mathrm{str}\,(x)$ hinaus ermittelt werden, so erweisen sich die dabei auftretenden vielstelligen Summenwerte als lästig, und es erscheint dann als wünschenswert, hierin eine Erleichterung herbeizuführen. Das ist in der Tat möglich, wie die im nachstehenden Abschnitt zu entwickelnde *zweite Form* der Summenmethode lehren wird.

Zweiundzwanzigste Vorlesung.

Numerische Bearbeitung: Summenmethode (II. Form).

190. Bei der Umformung der Summenmethode setzen wir wieder wie im vorigen Abschnitt einen unstetigen K.-G. voraus, dessen Argument x auf ganzzahlige Werte beschränkt ist, während m den Umfang und $\mathfrak{U}(x)$ die Verteilungsfunktion bedeutet; außerdem soll k und g das kleinste und das größte volle Argument sein. Dies festgesetzt greifen wir ein zwischen k und g gelegenes Argument $x = e$ heraus und denken uns von der Reihe der $m\mathfrak{U}(x)$ für den Augenblick nur die Glieder von $x = k$ bis $x = e$, d. h. also die Größen

$$m\mathfrak{U}(k), \quad m\mathfrak{U}(k+1), \cdots m\mathfrak{U}(e) \tag{1}$$

beibehalten, deren Anzahl offenbar durch den Ausdruck

$$H = e - k + 1 \tag{2}$$

gegeben ist. Die beibehaltenen Glieder behandeln wir nun weiter in in derselben Weise, wie früher die Größen $a(1), a(2), \ldots a(n)$, wobei H an die Stelle von n tritt. Demgemäß ist, wenn X wieder das Nummernargument bedeutet,

$$X = x - k + 1,$$
$$a(X) = m\mathfrak{U}(x) = m\mathfrak{U}(X + k - 1).$$

Ferner wird für die aus den Summengrößen $(H+1, p+1)$ folgenden Paarsummen nach § 186 (16)

$$(H+1,\, 1)_q = \sum_Z a(H-Z)Z^q = \sum_Z m\mathfrak{U}(e-Z)Z^q,$$

woraus mit $e - Z = x$

$$(H+1,\, 1)_q = \sum_x m\mathfrak{U}(x)(e-x)^q, \quad (x = k, \cdots e) \tag{3}$$

folgt.

Nachdem $(H+1,\, 1)_q$ gefunden ist, wiederholen wir, zu der ursprünglichen Reihe der $m\mathfrak{U}(x)$ zurückkehrend, die ganze Rechnung mit der Abänderung, daß jetzt nur die Glieder von $x = e$ bis $x = g$ und zwar in umgekehrter Reihenfolge benutzt werden. Es sind dann die Werte von $a(1), \ldots$ durch die Größen

$$m\mathfrak{U}(g), \quad m\mathfrak{U}(g-1), \cdots m\mathfrak{U}(e) \tag{4}$$

gegeben. Dementsprechend setzen wir an

$$J = g - e + 1, \qquad X = g - x + 1, \tag{5}$$

$$a(X) = m\mathfrak{U}(x) = m\mathfrak{U}(g - X + 1), \tag{6}$$

$$(J+1,\, 1)_q = \sum_Z a(J-Z)Z^q = \sum_Z m\mathfrak{U}(Z+e)Z^q,$$

woraus mit $Z + e = x$

$$(-1)^q (J+1,\,1)_q = \sum_x m\mathfrak{U}(x)(e-x)^q, \quad (x = e, \cdots g) \tag{7}$$

folgt. Vergleicht man nun die Summen auf den rechten Seiten von (3) und (7), so ist der zu summierende Ausdruck beidemal derselbe, ferner kommen mit Ausnahme von $x = e$ die Argumente von $x = k$ bis $x = g$ nur je einmal vor. Setzt man also zur Abkürzung

$$A(q) = m\mathfrak{D}[(e-x)^q],$$

so wird

$$A(q) = (H+1,\,1)_q + (-1)^q (J+1,\,1)_q - m\mathfrak{U}(e)(e-e)^q, \tag{8}$$

wo die Potenz $(e-e)^q$ für $q = 0$ gleich Eins, sonst aber gleich Null zu setzen ist.

§ 191. Aus den $A(q)$ hätte man jetzt die D-Koeffizienten herzuleiten. Um hierbei eine Kontrolle zu schaffen, die kürzer und auch sicherer ist, als es die einfache Wiederholung der Rechnung sein würde, schlagen wir folgenden Weg ein. Man denke sich die Operation, aus der (8) entstanden ist, nochmals ausgeführt, jedoch mit der Abänderung, daß man den Schnitt zwischen den $m\mathfrak{U}(x)$ nicht an der Stelle $x = e$, sondern an der nächstfolgenden Stelle $x = f = e + 1$ vornimmt. Setzt man dann

$$K = f - k + 1, \quad L = g - f + 1, \tag{9}$$

wobei K und L an die Stelle von H und J treten, so erhält man für den Ausdruck

$$B(q) = m\mathfrak{D}[(f-x)^q]$$

die Darstellung

$$B(q) = (K+1,\,1)_q + (-1)^q (L+1,\,1)_q - m\mathfrak{U}(f)(f-f)^q. \tag{10}$$

Nun läuft $m\mathfrak{U}(x)$ bei der Rechnung

$$\begin{array}{lllll}
\text{für } H & \text{von} & m\mathfrak{U}(k) & \text{bis} & m\mathfrak{U}(e) = m\mathfrak{U}(f-1), \\
\;\;,, \;\; K & ,, & m\mathfrak{U}(k) & ,, & m\mathfrak{U}(f), \\
\;\;,, \;\; J & ,, & m\mathfrak{U}(g) & ,, & m\mathfrak{U}(e), \\
\;\;,, \;\; L & ,, & m\mathfrak{U}(g) & ,, & m\mathfrak{U}(f) = m\mathfrak{U}(e+1),
\end{array}$$

so daß die Summenreihen zu K und J die Reihen zu H und L bereits in sich schließen; man findet, wenn man nach der Tabelle II in § 184 mit Feldern rechnet, die Werte zu H und L aus den vorletzten Feldern der Rechnung zu K und J. Demnach besteht die Mehrarbeit nur in der Berechnung der Paarsummen zu K und L. Gleichzeitig erhält man aber auch mit einer kleinen Nebenrechnung eine wirksame Kontrolle der $A(q)$ und $B(q)$. Man schreibe nämlich

$$A(0), \quad A(1), \quad A(2), \quad \ldots$$

hin und addiere je zwei benachbarte Glieder, darauf wiederhole man diesen Prozeß an den entstandenen Paarsummen und fahre damit fort, bis die Rechnung von selber abbricht. Dann sind, wie man bei Beachtung der in § 189 entwickelten Beziehungen unschwer erkennt, die Anfangsglieder der einzelnen Paarsummenreihen nichts anderes als die $B(q)$.

Bei der Anwendung von (8) und (10) ist es in der Regel vorteilhaft, das Argument e so zu wählen, daß der Argumentdurchschnitt $\mathfrak{D}(x)$ in die Strecke von e bis f fällt. Man hat dann nämlich bei dem Übergange auf die Normalform immer nur mit kleinen Reduktionen zu tun und kommt, falls man logarithmisch rechnet, mit vier Dezimalen aus. Aus diesem Grunde empfiehlt es sich, bevor man nach (8) und (10) zu rechnen beginnt, zunächst $\mathfrak{D}(x)$ für sich zu bestimmen. Hierzu genügt es, nach dem in § 187 entwickelten Verfahren die beiden niedrigsten Summenreihen zu der unzerschnittenen Reihe der $m\mathfrak{U}(x)$ zu rechnen.

§ 192. Als numerisches Beispiel soll jetzt wieder die in § 188 als Tabelle III mitgeteilte Verteilung der Endnullen in dem „Thesaurus" von *Vega* dienen. Unter der Voraussetzung, daß $\mathfrak{D}(x)$ in der soeben angegebenen Weise oder auch auf irgend eine andere Art als zwischen $x = 6$ und $x = 7$ liegend ermittelt worden sei, setzen wir zunächst an

$$k = 1, \quad g = 14, \quad e = 6, \quad f = 7,$$
$$H = 6, \quad J = 9, \quad K = 7, \quad L = 8.$$

Die weitere Rechnung ist in den einzelnen jetzt folgenden Tabellen enthalten, und zwar bis $q = 6$ hin, indem wir uns die Aufgabe stellen, die D-Koeffizienten bis zur Ordnung 6 zu ermitteln.

In Tabelle IV und V ist die Berechnung der Summenreihen zu H, J, K, L bis zur 7-ten Ordnung enthalten, und zwar zur besseren Ausnutzung des Raumes in derselben Anordnung, die oben in § 184 für Tabelle I benutzt worden war. Die darauf folgenden Tabellen VI und VII enthalten die Berechnung der Paarsummen und bedürfen einer näheren Erläuterung. Die vorletzte Spalte von Tabelle IV enthält, wenn man die in der ersten Zeile stehende a-Größe beiseite läßt, die Summengrößen $(H + 1, p + 1)$ oder ausgeschrieben

$$(H + 1, 1), \quad (H + 1, 2), \quad \ldots \quad (H + 1, 7),$$

d. h. die 7 Zahlen

$$613, \quad 867, \quad 659, \quad 282, \quad 66, \quad 6, \quad 0, \tag{11}$$

die mit 0, 1, ... multipliziert die Produkte

$$0, \quad 867, \quad 1318, \quad 846, \quad 264, \quad 30, \quad 0 \tag{12}$$

liefern, aus denen die 6 Paarsummen $(H + 1, p + 1)_1$

$$867, \quad 2185, \quad 2164, \quad 1110, \quad 294, \quad 30 \tag{13}$$

entspringen. Multipliziert man in der letzten Reihe wieder mit 0, 1, ..., so entsteht die Reihe

$$0, \quad 2185, \quad 4328, \quad 3330, \quad 1176, \quad 150, \tag{14}$$

mit den 5 Paarsummen $(H+1, p+1)_2$

$$2185, \quad 6513, \quad 7658, \quad 4506, \quad 1326. \tag{15}$$

Die dritte Operation liefert

$$0, \quad 6513, \quad 15316, \quad 13518, \quad 5304 \tag{16}$$

mit den 4 Paarsummen $(H+1, p+1)_3$

$$6513, \quad 21829, \quad 28834, \quad 18822. \tag{17}$$

Tabelle IV: Rechnung für H und K.

6	36	78	149	161	183	134	—
0	6	42	120	269	430	613	747
0	0	6	48	168	437	867	1480
0	0	0	6	54	222	659	1526
0	0	0	0	6	60	282	941
0	0	0	0	0	6	66	348
0	0	0	0	0	0	6	72
0	0	0	0	0	0	0	6

Tabelle V: Rechnung für J und L.

2	0	10	19	34	74	114	134	183	—
0	2	2	12	31	65	139	253	387	570
0	0	2	4	16	47	112	251	504	891
0	0	0	2	6	22	69	181	432	936
0	0	0	0	2	8	30	99	280	712
0	0	0	0	0	2	10	40	139	419
0	0	0	0	0	0	2	12	52	191
0	0	0	0	0	0	0	2	14	66

In dieser Weise geht es weiter. Vergleicht man nun die Reihen (11), (13), (15) und (17) miteinander, so erkennt man, daß das Anfangsglied jeder Paarsummenreihe in der vorhergehenden Reihe als zweites Glied auftritt, entsprechend der für die Bildung der Paarsummen geltenden Formel

$$(H+1, p+1)_{q+1} = p(H+1, p+1)_q + (p+1)(H+1, p+2)_q,$$

aus der für $p=0$

$$(H+1, 1)_1 = (H+1, 2)_0, \quad (H+1, 1)_2 = (H+1, 2)_1, \quad \ldots$$

folgt. Man darf also in jeder Paarsummenreihe das Anfangsglied und in jeder Produktreihe das Nullprodukt einfach fortlassen, ohne daß

man eine notwendige Zahl verliert. Dementsprechend ist das Schema von Tabelle VI und VII angeordnet. Die obere Hälfte von Tabelle VI enthält in der ersten Spalte die Werte aus (11) und zwar in um-

Tabelle VI: Berechnung von $A(q)$.

0	*0*				
6	*30*				
	30	*150*			
66	*264*				
	294	*1176*			
282	*846*	1326	*5 304*		
	1110	*3330*			
659	*1318*	4506	*13 518*		
	2164	*4328*	18 822	*56 466*	
867	*867*	7658	*15 316*		
	2185	*2185*	28 834	*57 668*	
613		**6513**	*6 513*	114 134	*228 268*
			21 829	*21 829*	
				79 497	*79 497*
					307 765
66	*396*				
191	*955*				
	1351	*6 755*			
419	*1676*				
	2631	*10 524*			
712	*2136*	17 279	*69 116*		
	3812	*11 436*			
936	*1872*	21 960	*65 880*		
	4008	*8 016*	134 996	*404 988*	
891	*891*	19 452	*38 904*		
	2763	*2 763*	104 784	*209 568*	
570		**10 779**	*10 779*	614 556	*1 229 112*
			49 683	*49 683*	
				259 251	*259 251*
					1 488 363
− 24	**+ 4948**	**− 4 266**	**+ 71 512**	**− 179 754**	**+ 1 796 128**

gekehrter Reihenfolge. Das zweite Glied von unten ist also $(H+1, 2)_0 = (H+1, 1)_1$. Neben den eingetragenen Zahlen sind in der zweiten Spalte (in kleinerer Schrift) die mit den Faktoren 6, 5, ... gebildeten Vielfachen der ersten Spalte angesetzt. Diese Produkte

summiert man paarweise von unten herauf und schreibt dabei jede Summe unmittelbar unter den unteren Summandus. Die unterste Zahl ist dann der Wert von $(H+1,\ 1)_2$. Weiter multipliziert man

Tabelle VII: Berechnung von $B(q)$.

6	36				
72	360				
	396	1 980			
348	1392				
	1752	7 008			
941	2823	8 988	35 952		
	4215	12 645			
1526	3052	19 653	58 959		
	5875	11 750	94 911	284 733	
1480	1480	24 395	48 790		
	4532	4 532	107 749	215 498	
747		16 282	16 282	500 231	1 000 462
			65 072	65 072	
				280 570	280 570
					1 281 032
14	84				
52	260				
	344	1720			
139	556				
	816	3264			
280	840	4984	19 936		
	1396	4188			
432	864	7452	22 356		
	1704	3408	42 292	126 876	
504	504	7596	15 192		
	1368	1368	37 548	75 096	
387		4776	4 776	201 972	403 944
			19 968	19 968	
				95 064	95 064
					499 008
+ 976	+ 5900	+ 11 506	+ 85 040	+ 185 506	+ 1 780 040

die Paarsummen der zweiten Spalte mit 5, 4, ..., setzt die Produkte in der dritten Spalte jedesmal auf die Zeilen des Multiplikandus, bildet die Paarsummen und so fort bis zum Schluß. Damit sind dann die H-Größen gefunden.

Die untere Hälfte von Tabelle VI enthält die Rechnung zu J, ferner ist in Tabelle VII die Rechnung zu K und L gegeben. Die Größen

$$(H+1,\,1)_q,\quad (J+1,\,1)_q,\quad (K+1,\,1)_q,\quad (L+1,\,1)_q,$$

die das eigentliche Ergebnis dieser Zwischenrechnung bilden, sind hierbei durch fetteren Druck hervorgehoben. Werden die vorstehenden Größen nach den Formeln (8) und (10) miteinander verbunden, so erhält man die Größen $A(q)$ und $B(q)$, deren Werte am Fuße jeder Tabelle angesetzt sind, jedoch unter Fortlassung von $A(0)$ und $B(0)$, deren Werte mit dem Umfange m identisch sind.

Tabelle VIII enthält die Kontrollrechnung für die $A(q)$ und $B(q)$. In die erste Spalte sind die A-Größen nach Tabelle VI eingetragen.

Tabelle VIII: Kontrolle der $A(q)$ und $B(q)$.

+ 1 000						
	+ 976					
− 24		+ 5 900				
	+ 4 924		+ 11 506			
+ 4 948		+ 5 606		+ 85 040		
	+ 682		+ 73 534		+ 185 506	
− 4 266		+ 67 928		+ 100 466		+ 1 780 040
	+ 67 246		+ 26 932		+ 1 594 534	
+ 71 512		− 40 996		+ 1 494 068		
	− 108 242		+ 1 467 136			
− 179 754		+ 1 508 132				
	+ 1 616 374					
+ 1 796 128						

Diese werden paarweise summiert; die Paarsummen stehen in der zweiten Spalte, die in derselben Weise behandelt wird, und so fort bis zur letzten Spalte hin. Die Anfangsglieder der einzelnen Spalten stimmen, wie es sein muß, mit den $B(q)$ in Tabelle VII überein.

Die Tabelle IX enthält endlich die Zusammenstellung der Zahlen, die als Ausgangspunkt für die weitere Rechnung zu dienen haben.

Tabelle IX.

q	$C(q)$ für $e = 6$			$C(q)$ für $e = 7$		
0	+	**1 000**	3.0000	+	**1 000**	3.0000
1	−	**24**	1.3802n	+	**976**	2.9894
2	+	**4 948**	3.6944	+	**5 900**	3.7709
3	−	**4 266**	3.6300n	+	**11 506**	4.0609
4	+	**71 512**	4.8544	+	**85 040**	4.9296
5	−	**179 754**	5.2547n	+	**185 506**	5.2684
6	+	**1 796 128**	6.2544	+	**1 780 040**	6.2504

Hierbei soll von jetzt ab e gemeinsam für e und f, ferner w für $e-x$, sowie

$$C(q) = m\mathfrak{D}(w^q) \qquad (18)$$

gemeinsam für $A(q)$ und $B(q)$ geschrieben werden. Dementsprechend enthält die Tabelle die beiden Wertreihen der $C(q)$ nebst den aus einer vierstelligen Tabelle entnommenen Logarithmen der $C(q)$.

Für die weitere Rechnung ist es gleichgültig, ob die $C(q)$ nach der Summenmethode oder durch Summierung der direkt gebildeten Produkte $m\mathfrak{U}(x)w^q$ oder sonstwie gefunden worden sind. Ferner will ich den an sich bequemen Umstand, daß der Umfang m eine runde Zahl ist, absichtlich außer acht lassen und demgemäß das Schema so anlegen, als ob m eine beliebige Zahl wäre.

§ 193. Der nächste Schritt besteht in der Ermittlung der Größen $\mathfrak{D}(x)$, $\operatorname{str}(x)$ und h. Zieht man in Tabelle IX die Logarithmen der ersten Zeile von den folgenden Logarithmen ab, so erhält man die Logarithmen der Größen $\mathfrak{D}(w^q)$. Damit stellt sich die Rechnung für

$$\operatorname{str}(x)^2 = \mathfrak{D}(w^2) - \mathfrak{D}(w)^2, \qquad h = 1 : \sqrt{2}\operatorname{str}(x)$$

folgendermaßen.

Tabelle X.

	$e = 6$	$e = 7$	Mittel
Log $\mathfrak{D}(w)$	8.3802 n	9.9894	
Log $\mathfrak{D}(w)^2$	6.7504	9.9788	
Log $\mathfrak{D}(w^2)$	0.6944	0.7709	
$\mathfrak{D}(e-x)$	$-$ 0.0240	$+$ 0.9759	
$\mathfrak{D}(x)$	$+$ 6.0240	$+$ 6.0241	$+$ 6.0240
$\mathfrak{D}(w^2)$	4.948	5.901	
$\mathfrak{D}(w)^2$	0.001	0.952	
$\operatorname{str}(x)^2$	4.947	4.949	4.948
Log $\operatorname{str}(x)^2$	—	—	0.6944
Log $\operatorname{str}(x)$	—	—	0.3427
Log $(1 : h)$	—	—	0.4977
Log h	—	—	9.5023
$\operatorname{str}(x)$	—	—	2.224
h	—	—	0.3179

Handelt es sich um einen stetigen K.-G., so ist der für $\operatorname{str}(x)$ gefundene Wert noch wegen der Abrundung zu verbessern. Die nötige Reduktion ergibt sich, wenn der Einfluß der Abrundungsphase vernachlässigt werden darf, und wenn ferner s den verbesserten Wert bedeutet, nach § 142 (47) aus der Gleichung

$$s^2 = \operatorname{str}(x)^2 - \frac{1}{3}T^2,$$

wo T die halbe Teilstreckenlänge bedeutet. Nun ist bei ganzzahligen Argumenten $2T = 1$, also wird die an $\operatorname{str}(x)^2$ anzubringende negative Verbesserung gleich $1 : 12$ oder gleich 0.0833. Damit würde sich

die Rechnung in der letzten Spalte der obigen Tabelle folgendermaßen gestalten:

Tabelle X.A

$\mathrm{str}(x)^2 =$	4.948	$\mathrm{Log}\, s^2$	$= 0.6870$	$\mathrm{Log}\, h$	$= 9.5060$
Verb. $=$	-0.083	$\mathrm{Log}\, s$	$= 0.3435$	s	$= 2.205$
$s^2 =$	4.865	$\mathrm{Log}\,(1:h)$	$= 0.4940$	h	$= 0.3206$

Die Reduktion beträgt in s und h rund ein Prozent, ist also keineswegs unmerklich, soll jedoch zunächst beiseite gelassen werden.

§ 194. Bei der Berechnung der übrigen numerischen Elemente wollen wir den Umstand berücksichtigen, daß die im Anhange mitgeteilten Tafeln für die Größen Φ_q nicht unmittelbar die Werte dieser Funktionen, sondern die Beträge von

$$\Psi_q = \Phi_q : 2^{q-1}$$

angeben. Man erreicht dadurch, daß die extremen Werte der tabulierten Funktionen in der Nähe der Einheit bleiben, was für manche Zwecke angenehm ist. Dementsprechend wird man bei der numerischen Rechnung die Produkte $D_q \Phi_q$ in der Form

$$D_q \Phi_q = [2^{q-1} D_q] \cdot \Psi_q \tag{19}$$

bilden, d. h. den Divisor in Ψ_q als Faktor zu D_q schlagen. Hiernach würde es sich empfehlen, von vornherein auf die Größen $2^{q-1} D_q$ auszugehen, wenn auch in den analytischen Entwicklungen die ursprüngliche Form $D_q \Phi_q$ für die Rechnung geschmeidiger ist. Noch bequemer ist es jedoch, auf den Zwischenstufen der Rechnung die Größen $2^q D_q$ zu benutzen, deren Halbierung dann die in (19) gebrauchten Koeffizienten der Ψ_q liefert.

Bezeichnen wir mit $D(e, k)_q$ die mit beliebig gewählten Parametern e und k berechneten D-Koeffizienten, so soll

$$E(e, k)_q = 2^q D(e, k)_q = 2^q \mathfrak{D}[\Re(kx - ke)_q] \tag{20}$$

sein. Daraus folgt

$$\mathfrak{D}[\exp(-2kxv + 2kev - v^2)] = \sum_q D(e, k)_q (2v)^q = \sum_q E(e, k)_q v^q. \tag{21}$$

Nehmen wir für k den normalen Wert h und schreiben w für $e - x$, so wird hiernach

$$\sum_q E(e, h)_q v^q = \exp(-v^2) \mathfrak{D}[\exp(2hvw)]. \tag{22}$$

Entwickelt man die $\mathfrak{D}$-Größe auf der rechten Seite in die Reihe

$$\mathfrak{D}[\exp(2hvw)] = P(0) + P(1)v + P(2)v^2 + \cdots,$$

so ist

$$q!\, P(q) = \mathfrak{D}[(2hw)^q],$$

woraus mit Berücksichtigung von (18)

$$mq!\,P(q) = (2h)^q C(q) \tag{23.a}$$

folgt. Sind die $P(q)$ gefunden, so erhält man die E-Größen wegen (22) aus

$$E(e,h)_q = \frac{P(q)}{0!} - \frac{P(q-2)}{1!} + \frac{P(q-4)}{2!} - \cdots. \tag{23.b}$$

Hieraus sind endlich noch die für den normalen Wert $e = c = \mathfrak{D}(x)$ geltenden $E(c,h)_q$ herzuleiten. Zu dem Ende setzen wir $f = 2h(c-e)$ und schreiben (21) mit $k = h$ in der Gestalt

$$\mathfrak{D}[\exp(-2hxv + 2hcv - v^2)] = \exp(2hcv - 2hev)\sum_q E(e,h)_q v^q,$$

$$\sum_q E(c,h)_q v^q = \exp(fv)\sum E(e,h)_q v^q,$$

woraus

$$E(c,h)_q = E(e,h)_q + \frac{f}{1!}E(e,h)_{q-1} + \frac{f^2}{2!}E(e,h)_{q-1} + \cdots \tag{24.a}$$

folgt. Die Größe f braucht nicht besonders gerechnet zu werden, denn wegen der Eigenschaften der Normalform ist

$$E(c,h)_1 = E(c,h)_2 = 0,$$

so daß mit $q = 1$ aus (24.a) die Gleichungen

$$0 = E(e,h)_1 + f, \qquad f = -E(e,h)_1 \tag{24.b}$$

folgen.

Da die Beziehungen zwischen den Größen P und E homogen linear sind, so kann man auch mit den Größen $mP(q)$ und mE_q rechnen und entsprechend die Größen mD_q ansetzen. Das soll nachstehend geschehen, und zwar aus folgendem Grunde. Die beobachteten $m\mathfrak{U}(x)$ sind ihrer Entstehung nach ganze Zahlen, so daß die Einheit bei ihnen die natürliche Grenze für die Schärfe der Notierung bildet. Demgemäß wird man bei der Ansetzung der theoretisch gefundenen Werte von $m\mathfrak{U}(x)$ nicht über die erste Dezimale hinauszugehen haben, und das Gleiche gilt von den Größen $mP(q)$ und mE_q. Es stellt sich also, wenn man mit den $mP(q)$ und mE_q rechnet und dabei nur die erste Dezimale mitnimmt, ganz von selbst diejenige Schärfe der Rechnung ein, die dem jedesmaligen Betrage von m angemessen ist, wenn man unnötigen Ziffernballast vermeiden will.

§ 195. Es soll nnn das oben angefangene Beispiel an der Hand der entwickelten Formeln weiter verfolgt werden. Die Rechnung dazu ist in Tabelle XI gegeben. Die Spalten sind mit den Überschriften (0), (1), ... versehen, die den Indexwerten $q = 0, 1, \ldots$ entsprechen. Linkerhand sind zur leichteren Beschreibung die Zeilennummern angesetzt, dagegen sind des Raumes wegen die, übrigens leicht zu ergänzenden, Angaben für die Bedeutung der Zeilen fortgelassen.

Tabelle XI.

	(0)	(1)	(2)	(3)	(4)	(5)	(6)
1	0.0000	0.0000	9.6990	9.2218	8.6198	7.9208	7.1427
2	0.0000	9.8033	9.6066	9.4099	9.2132	9.0165	8.8198
3	3.0000	1.3802 *n*	3.6944	3.6300 *n*	4.8544	5.2547 *n*	6.2544
4	0.0000	9.8033	9.3056	8.6317	7.8330	6.9373	5.9625
5	3.0000	2.9894	3.7709	4.0609	4.9296	5.2684	6.2504
6	3.0000	1.1835 *n*	3,0000	2.2617 *n*	2.6874	2.1920 *n*	2.2169
7	+ 1000.0	— 15.3	+ 1000.0	— 182.7	+ 486.8	— 155.6	+ 164.8
8			— 1000.0	+ 15.3	— 1000.0	+ 182.7	— 486.8
9					+ 500.0	— 7.6	+ 500 0
10							— 166.7
11	+ 1000.0	— 15.3	0.0	— 167.4	— 13.2	+ 19.5	+ 11.3
12		+ 15.3	— 0.2	0.0	— 2.6	— 0.2	+ 0.3
13			+ 0.1	0.0	0.0	0.0	0.0
14				0.0	0.0	0.0	0.0
15					0.0	0.0	0.0
16						0.0	0.0
17							0.0
18	+ 1000.0	0.0	— 0.1	— 167.4	— 15.8	+ 19.3	+ 11.6
19	3.0000	1.1835 *n*	—	2.2237 *n*	1.1206 *n*	1.2900	—
20		1.1835	9.3670 *n*	—	0.4072 *n*	9.3041 *n*	9.4735
21			9.0660	—	—	—	—
22				—	—	—	—
23					—	—	—
24						—	—
25							—
26	3.0000	2.7927	3.0765	2.6926	2.7626	2.2057	2.2129
27	+ 1000.0	+ 620.4	+ 1192.7	+ 492.7	+ 578.9	+ 160.6	+ 163.3
28			— 1000.0	— 620.4	— 1192.7	— 492.7	— 578.9
29					+ 500.0	+ 310.2	+ 596.3
30							— 166.7
31	+ 1000.0	+ 620.4	+ 192.7	— 127.7	— 113.8	— 21.9	+ 14.0
32		— 620.4	— 385.0	— 119.6	+ 79.2	+ 70.6	+ 13.6
33			+ 192.5	+ 119.4	+ 37.1	— 24.6	— 21.9
34				— 39.8	— 24.7	— 7.7	+ 5.1
35					+ 6.2	+ 3.8	+ 1.2
36						— 0.8	— 0.5
37							+ 0.1
38	+ 1000.0	0.0	+ 0.2	— 167.7	— 16.0	+ 19.4	+ 11.6
39	3.0000	2.7927	2.2850	2.1062 *n*	2.0561 *n*	1.3404 *n*	—
40		2.7927 *n*	2.5854 *n*	2.0777 *n*	1.8989	1.8488	1.1331
41			2.2844	2.0771	1.5694	1.3906 *n*	1.3407 *n*
42				1.6000 *n*	1.3927 *n*	0.8850 *n*	0.7062
43					0.7906	0.5833	0.0756
44						9.8843 *n*	9.6770 *n*
45							8.8988

Zeile 1 enhält die Logarithmen von $1 : q!$, Zeile 2 die Logarithmen von $(2h)^q$. Beide Zeilen werden summiert und die Summen in Zeile 4 eingetragen, die hiernach die Logarithmen von $(2h)^q : q!$ angibt. Ferner werden die in Tabelle IX aufgeführten Logarithmen der $C(q)$ eingetragen, und zwar das erste der beiden C-Systeme in Zeile 3, das zweite in Zeile 5. Darauf werden die Summen von Zeile 3 und 4 in Zeile 6, die Summen von Zeile 4 und 5 in Zeile 26 gesetzt. Die Zeilen 6 und 26 enthalten dann die Logarithmen der Größen $mP(q)$, die zu den beiden benutzten Werten von e gehören. Die an Zeile 6 anschließende Rechnung steht in den Zeilen 7 bis 25; die entsprechende als Kontrolle dienende zweite Rechnung schließt sich in derselben Weise an Zeile 26 an.

Zeile 7 enthält die Numeri von Zeile 6 oder die $mP(q)$ der ersten Rechnung. Aus diesen Größen sind nun nach Formel (23.b) die $mE(e, h)$ abzuleiten. Zu dem Ende multipliziert man die Zeile 7 mit $-1 : 1$ und schreibt die Produkte in Zeile 8, aber um zwei Spalten nach rechts verschoben. Dann multipliziert man Zeile 8 mit $-1 : 2$ und setzt die Produkte, wiederum um zwei Spalten verschoben, in Zeile 9. Ebenso multipliziert man Zeile 9 mit $-1 : 3$ und setzt die Produkte um zwei Spalten verschoben in Zeile 10. Damit sind die nach der Formel (23.b) nötigen Glieder gefunden, weil wir hier nur bis zu dem Index $q = 6$ gehen wollen. Die in Zeile 11 angesetzten Summen der Zeilen 7 bis 10 liefern die $mE(e, h)_q$ der ersten Rechnung. In derselben Weise entstehen die Zeilen 27 bis 31 aus der Zeile 26.

Die gefundenen Größen $mE(e, h)_q$, die zu den beiden Werten von e gehören, müssen nun auf die Normalform reduziert werden, zu welchem Ende die verschiedenen Glieder der Formel (24.a) zu bilden sind. Bei der Beschreibung dieser Operation möge die zweite Rechnung zugrunde gelegt werden, weil sie die größeren Zahlen enthält. Die nach Formel (24.a) zu bildenden Glieder sind in den Zeilen 31 bis 37 enthalten, während die bei der Zwischenrechnung auftretenden Logarithmen jedesmal 8 Zeilen tiefer unter den Zeilennummern 39 bis 45 stehen. Zunächst setzt man in 39 die Logarithmen von 31 an, wobei in Spalte (6) ein Strich steht, weil der betreffende Logarithmus nicht gebraucht wird. Auch sind die Logarithmen der in den Spalten (0) und (1) stehenden Zahlen nicht erst aufzuschlagen, weil sie schon auf Zeile 26 vorkommen. Des weiteren kommt den beiden ersten Werten der Zeile 39 folgende Bedeutung zu:

$$\text{in Spalte (0):}\quad \mathrm{Log}\,[mE(e, h)_0] = \mathrm{Log}\, m,$$

$$\text{in Spalte (1):}\quad \mathrm{Log}\,[mE(e, h)_1]\,,$$

woraus nach Formel (24.b)

$$\mathrm{Log}\,(-f) = \text{Spalte (1)} - \text{Spalte (0)} = 9.7927$$

also $\operatorname{Log} f = 9.9727n$ folgt. Diesen Wert von $\operatorname{Log} f$ addiert man zu den Werten der Zeile 39 und trägt die jedesmal erhaltene Summe eine Spalte weiter rechts in Zeile 40 ein. Dann addiert man zu den Werten der Zeile 40 den Logarithmus von $f:2$ und trägt die Summen eine Spalte weiter rechts in Zeile 41 ein. Ebenso entsteht die Zeile 42 aus 41 mit dem Logarithmus von $f:3$ usw., bis die letzte Zahl in Spalte (6) Zeile 45 gefunden ist. Darauf werden die Numeri der so erhaltenen Logarithmen 8 Zeilen höher in den Zeilen 32 bis 37 eingetragen; die Summation der Zeilen 31 bis 37 liefert dann endlich die gesuchten normalen Werte der Größen $mE(c,h)_q$ in Zeile 38.

Geht man nunmehr die nach demselben Schema verlaufende erste Rechnung in den Zeilen 11 bis 25 durch, so wird zunächst für $\operatorname{Log} f$ der Betrag 8.1835 gefunden. Da dieser Wert ziemlich klein ist, so liegt die Mehrzahl der Logarithmen, die in den Zeilen 20 bis 25 einzutragen wären, unterhalb des Betrages 8.699 oder Log(0.05). Die entsprechenden in den Zeilen 12 bis 17 einzutragenden Numeri würden daher unter 0.05 liegen, also bei der festgehaltenen Rechnungsschärfe einfach durch Null zu ersetzen sein. Da es hiernach keinen Zweck hätte, die Logarithmen der unmerklichen Reduktionsgrößen erst zu bilden, so ist an den betreffenden Stellen der Logarithmus durch einen Strich ersetzt worden.

Die in den Zeilen 18 und 38 stehenden Zahlen sind die durch doppelte Rechnung gefundenen Normalwerte der Größen $mE(c,h)_q$, die wir jetzt kurz durch mE_q bezeichnen wollen. Da die beiden Zahlen soweit übereinstimmen, als mit Rücksicht auf die benutzte Stellenzahl zu erwarten ist, und da ferner die Beträge der Größen mE_1 und mE_2, die bei völlig scharfer Rechnung verschwinden müssen, von ihren Sollwerten nur unmerklich abweichen, so können wir jetzt die beiden Zeilen zu einem Mittel zusammenziehen. Hierbei lassen wir mE_0, das ja mit dem Umfange m identisch ist, beiseite; ebenso dürfen wir für mE_1 und mE_2 ihre Sollwerte setzen, so daß nur die Größen E_3, E_4, E_5, E_6 in Betracht kommen. Tabelle XII enthält diese Zahlen übersichtlich zusammengestellt.

Tabelle XII.

q	3	4	5	6
mE_q: 1. Rechn.	− 167.4	− 15.8	+ 19.3	+ 11.6
mE_q: 2. Rechn.	− 167.7	− 16.0	+ 19.4	+ 11.6
$2mE_q$	− 335.1	− 31.8	+ 38.7	+ 23.2
F_q	− 41.9	− 4.0	+ 4.8	+ 2.9
E_q	− 0.1675	− 0.0159	+ 0.0193	+ 0.0116

Die zweite und dritte Zeile enthält die beiden Wertsysteme der mE_q. Darunter ist die Summe der beiden Zeilen angesetzt, aus der durch

Division mit $2m$ die in der letzten Zeile aufgeführten Werte der E_q folgen. Außerdem sind in der vorletzten Zeile noch die weiterhin gebrauchten Werte der Größe $F_q = mE_q : 4$ angesetzt, die aus der Zeile darüber durch Division mit 8 folgen. Handelt es sich darum, für eine größere Anzahl von Kollektivreihen die Elemente übersichtlich zusammenzustellen, so empfiehlt es sich, neben den Größen m, $\mathfrak{D}(x)$, $\operatorname{str}(x)$, h nicht die D_q, sondern die E_q für $q = 3, 4, \ldots$ aufzuführen, weil man dann Dezimalen spart.

§ 196. Wenn die hier untersuchten Verteilungszahlen einem stetigen K.-G. angehörten, so würden die in Tabelle XI und XII auftretenden Größen noch wegen der Abrundung zu reduzieren sein. Wir wollen nun noch zusehen, wie sich die Rechnung gestaltet, wenn man annimmt, daß die beobachteten Zahlen sich auf einen stetigen K.-G. beziehen. Voraussetzung ist dabei, daß der Einfluß der Abrundungsphase vernachlässigt werden dürfe, da man andernfalls entweder auf die Elimination der Abrundung oder aber auf das Verfahren der direkten Berechnung der D-Koeffizienten verzichten müßte. Da nun die mit den Ordinaten $m\mathfrak{U}(x)$ konstruierte Punktreihe sich nicht erheblich von der Glockenform des E.-G. entfernt, und da immerhin 14 Ordinaten vorliegen, so dürfen wir nach den Entwicklungen in der XVI. Vorlesung annehmen, daß der Einfluß der Abrundungsphase hinreichend klein ist, daß also nur der Einfluß der Teilstreckenlänge in Betracht kommt.

Die Stelle, wo man die Verbesserung in den Rechnungsgang einschaltet, ist — rein theoretisch betrachtet — natürlich gleichgültig: man kann die erforderliche Reduktion nach Belieben an die $P(q)$ oder an die $E(e, h)_q$ oder endlich an die $E(c, h)_q$ anbringen. Wünscht man die Doppelrechnung möglichst lange beizubehalten, so empfiehlt es sich, die Reduktion schon bei den $P(q)$ oder den $E(e, h)_q$ vorzunehmen. Arbeitet man mit einer Rechenmaschine, was bei solchen Rechnungen keineswegs unzweckmäßig ist, so macht es praktisch keinen Unterschied, ob man die eine oder die andere dieser Größenreihen verbessert. Bei der logarithmischen Rechnung dagegen gewährt es eine kleine Ziffernersparnis, wenn die Reduktion schon bei den $P(q)$ erfolgt. Wir wollen uns daher die Formeln für diesen letzten Fall zurechtlegen.

Aus § 142 nehmen wir zunächst die Formeln (43.b) und (41), nämlich die Gleichungen

$$2v\Delta[\exp(-2uv)] = L\mathfrak{D}[\exp(-2uv)], \tag{25}$$

$$2v : L = 1 - a_1(2Jv)^2 + a_2(2Jv)^4 \tag{26}$$

herüber, deren Bestandteile folgende Bedeutung haben. Das Symbol Δ bezeichnet den Durchschnitt, wie er bei direkter Mittelbildung auf Grund der vorgelegten abgerundeten Argumente gefunden wird, wogegen das Zeichen $\mathfrak{D}$ den von Abrundung befreiten Durchschnitt

angibt, den man in Wirklichkeit jetzt sucht. Die Größe v ist eine willkürliche Veränderliche, u dagegen das Hilfsargument $h(x-c)$, dessen Parameter beliebig gewählt werden dürfen. Die Größe J ist gleich hT, wo T die halbe Teilstreckenlänge des Arguments x bedeutet. Endlich bezeichnen die Koeffizienten a gewisse rationale Zahlen, deren Werte a. a. O. zusammengestellt sind.

Für das in u enthaltene c setzen wir jetzt die beiden Zahlen e, die den in Tabelle IX aufgeführten $C(q)$ zugrunde liegen. Für h wählen wir den normalen, wegen Abrundung bereits verbesserten Wert; in Tabelle X.A war für $\operatorname{Log} h$ der Betrag 9.5060 gefunden worden, während der unverbesserte Wert 9.5023 gelautet hatte. Setzt man wieder $e-x=w$, so ist die Gleichung (18) jetzt in der Gestalt

$$C(q)=m\Delta(w^q)$$

zu schreiben, woraus nach (23.a)

$$mq!\,P(q)=(2h)^q C(q)=m\Delta[(2hw)^q] \tag{27}$$

folgt. Bedeutet ferner $Q(q)$ den verbesserten Wert von $P(q)$, so wird

$$mq!\,Q(q)=m\mathfrak{D}[(2hw)^q]. \tag{28}$$

Da $u=h(x-e)=-hw$ ist, so liefern die beiden Seiten von (25) die Reihen

$$2v\Delta[\exp(2hwv)]=2v\sum_q P(q)v^q,$$

$$L\mathfrak{D}[\exp(2hwv)]=L\sum_q Q(q)v^q,$$

womit

$$\sum_q Q(q)v^q=[1-a_1(2Jv)^2+\cdots]\sum_q P(q)v^q$$

erhalten wird. Daraus ergibt sich durch Spaltung nach v die gesuchte Formel

$$Q(q)=P(q)-a_1(2hT)^2P(q-2)+a_2(2hT)^4P(q-4)-\cdots, \tag{29}$$

in der die Zahlenfaktoren $a_1, a_2, \ldots$ aus § 142 zu entnehmen sind. Aus (29) ergeben sich die verbesserten $E(e,h)_q$, indem man für die weitere Rechnung die Formel (23.b) durch

$$E(e,h)_q=\frac{Q(q)}{0!}-\frac{Q(q-2)}{1!}-\frac{Q(q-4)}{2!}-\cdots \tag{30}$$

ersetzt, während der Übergang auf die Normalform wie früher nach (24) zu erfolgen hat. Vergleicht man (29) und (30), so erkennt man, daß beidemal derselbe Rechnungsgang vorliegt, nur daß bei (30) die Multiplikatoren

$$-1:1!,\quad +1:2!,\quad -1:3!,\quad +1:4!,\quad \ldots \tag{31}$$

auftreten, wogegen bei (29) laut § 142 die Multiplikatoren

$$-\frac{(2hT)^2}{6},\quad +\frac{7(2hT)^4}{360},\quad -\frac{31(2hT)^6}{15120},\quad +\frac{127(2hT)^8}{604800},\quad \cdots \tag{32}$$

benutzt werden. Bei der Ausführung der in (30) vorgeschriebenen Rechnung hatten wir in Tabelle XI die Multiplikation mit den einzelnen Faktoren (31) in sukzessive Multiplikationen mit den Faktoren

$$-1:1, \quad -1:2, \quad -1:3, \quad -1:4, \quad \cdots \tag{33}$$

zerlegt. In ähnlicher Weise wird man (32) durch die sukzessive Multiplikation mit den Faktoren

$$-\frac{(2hT)^2}{6}, \quad -\frac{7(2hT)^2}{60}, \quad -\frac{31(2hT)^2}{294}, \quad -\frac{127(2hT)^2}{1240}, \quad \cdots \tag{34}$$

ersetzen.

§ 197. Die Tabelle XIII enthält die ähnlich wie Tabelle XI angelegte Rechnung für den Fall der Stetigkeit der betrachteten Kollektivreihe, jedoch unter Fortlassung der Operationen, die in Tabelle XI auf den Zeilen 8 bis 25 und 28 bis 45 auszuführen waren, da

Tabelle XIII.

	(0)	(1)	(2)	(3)	(4)	(5)	(6)
1	0.0000	0.0000	9.6990	9.2218	8.6198	7.9208	7.1427
2	0.0000	9.8070	9.6140	9.4210	9.2280	9.0350	8.8420
3	3.0000	1.3802*n*	3.6944	3.6300*n*	4.8544	5.2547*n*	6.2544
4	0.0000	9.8070	9.3130	8.6428	7.8478	6.9558	5.9847
5	3.0000	2.9894	3.7709	4.0609	4.9296	5.2684	6.2504
6	3.0000	1.1872*n*	3.0074	2.2728*n*	2.7022	2.2105*n*	2.2391
6a			1.2338*n*	9.4210	1.2412*n*	0.5066	0.9360*n*
6b					9.3127	—	9.3201
6c							—
6d	+1000.0	− 15.4	+1017.2	− 187.4	+ 503.7	− 162.4	+ 173.4
6e			− 17.1	+ 0.3	− 17.4	+ 3.2	− 8.6
6f					+ 0.2	—	+ 0.2
6g							—
7	+1000.0	− 15.4	+1000.1	− 187.1	+ 486.5	− 159.2	+ 165.0
26	3.0000	2.7964	3.0839	2.7037	2.7774	2.2242	2.2351
26a			1.2338*n*	1.0302*n*	1.3177*n*	0.9375*n*	1.0112*n*
26b					9.3127	9.1091	9.3966
26c							—
26d	+1000.0	+ 625.8	+1213.1	+ 505.5	+ 599.0	+ 167.6	+ 171.8
26e			− 17.1	− 10.7	− 20.8	− 8.7	− 10.3
26f					+ 0.2	+ 0.1	+ 0.2
26g							—
27	+1000.0	+ 625.8	+1196.0	+ 494.8	+ 578.4	+ 159.0	+ 161.7

diese Operationen genau wie früher verlaufen. Ferner sind die gegen früher neu hinzukommenden Zeilen zur bequemeren Vergleichung mit Tabelle XI durch die Nummern 6a bis 6g und 26a bis 26g ge-

kennzeichnet. Zeile 1 und 2 enthalten wiederum die Logarithmen von $1:q!$ und $(2h)^q$, wobei für h sogleich der wegen Abrundung verbesserte Wert zu benutzen ist. Die Summe der Zeilen 1 und 2 steht wieder in Zeile 4, wogegen in Zeile 3 und 5 die aus Tabelle IX zu entnehmenden Wertreihen von Log $C(q)$ eingetragen sind. Die Summe von Zeile 3 und 4 kommt nach Zeile 6, die Summe von Zeile 4 und 5 dagegen nach Zeile 26. Die Zeilen 6 und 26 enthalten also die Logarithmen der $mP(q)$, die nunmehr noch auf die $mQ(q)$ zu reduzieren sind. Zu den Werten in Zeile 6 wird der Logarithmus des ersten Gliedes der Reihe (34) addiert, wobei im vorliegenden Falle $2T=1$ zu setzen ist. Die Summen kommen um zwei Spalten nach rechts verschoben in Zeile 6a. Dann addiert man zu Zeile 6a den Logarithmus des zweiten Gliedes in (34) und setzt die Summen um zwei Spalten verschoben in Zeile 6b; in derselben Weise leitet man mit dem dritten Gliede in (34) das Glied der Zeile 6c ab, das im vorliegenden Falle wegen seiner Unmerklichkeit durch einen bloßen Strich angedeutet ist. Die Numeri der Zeilen 6 bis 6c stehen vier Zeilen tiefer in 6d bis 6g; die Summen dieser Numeri liefern die in Zeile 7 angesetzten Werte der $Q(q)$. Aus letzteren wird dann genau wie in Tabelle XI das System der $E(e, h)_q$ und $E(c, h)_q$ gebildet. Die Kontrollrechnung schließt sich in derselben Weise an die Zeile 26 an und braucht nicht mehr besonders besprochen zu werden.

Die Durchführung der Rechnung liefert für die normalen E_q die beiden Wertreihen

mE_3	mE_4	mE_5	mE_6
-171.7	-16.2	$+20.0$	$+12.1$
-171.8	-16.4	$+20.3$	$+12.1$

die befriedigend miteinander stimmen.

Stellt man die Zahlen für die unstetige und die stetige Rechnung übersichtlich zusammen, so ergibt sich

	m	$\mathfrak{D}(x)$	str (x)	Log h
unstetig	1000	6.024	2.224	9.5023
stetig	1000	6.024	2.205	9.5060
	E_3	E_4	E_5	E_6
unstetig	-0.1675	-0.0159	$+0.0193$	$+0.0116$
stetig	-0.1717	-0.0163	$+0.0201$	$+0.0121$
	F_3	F_4	F_5	F_6
unstetig	-41.9	-4.0	$+4.8$	$+2.9$
stetig	-42.9	-4.1	$+5.0$	$+3.0$

Die Unterschiede zwischen den beiden Fällen sind, da keine übertriebene Abrundung vorliegt, nicht sehr groß, immerhin jedoch in str (x), h und E_3 für die hier innegehaltene Rechnungsschärfe sehr merklich.

§ 198. Nachdem die Koeffizienten bis D_6 ermittelt worden sind, handelt es sich noch darum festzustellen, wie weit im vorliegenden Falle die mitgenommenen Glieder der Φ-Reihe den beobachteten Verlauf der Verteilung darstellen. Zu dem Ende setzen wir die vorgelegte Kollektivreihe zunächst wieder als unstetig voraus. Ferner führen wir das normale Hilfsargument $u = h(x - c)$ ein und setzen

$$mE_q = 4F_q, \quad 2\,\Phi(u)_q = 2^q\,\Psi(u)_q, \tag{35}$$

wo die F_q aus Tabelle XII in § 195 zu entnehmen sind, und die Ψ_q, wie bereits in § 194 bemerkt wurde, in den Tafeln des Anhanges gefunden werden. Dies vorausgeschickt schreiben wir die Grundgleichung

$$2\mathfrak{S}(x) - 1 = \Phi(u) + D_3\Phi(u)_3 + D_4\Phi(u)_4 + \cdots$$

in der Gestalt

$$m\mathfrak{S}(x) - \tfrac{1}{2}m = \tfrac{1}{2}m\,\Phi(u) + F_3\,\Psi(u)_3 + F_4\,\Psi(u)_4 + \cdots, \tag{36}$$

wobei zu beachten ist, daß in (36) für die Veränderliche x nicht die Argumentwerte der Verteilungstafel, sondern die zwischen den Argumenten liegenden Halbierungspunkte zu setzen sind, welche hier die Rolle der Wechselpunkte stetiger Kollektivreihen zu übernehmen haben.

In den Tabellen XIV findet man die erforderliche Rechnung zusammengestellt, die des Raumes wegen in mehrere nebeneinander zu setzende Abschnitte zerlegt worden ist. Die Spalte x in XIV A

Tabelle XIVA.

x	u	$\Phi(u)$	$\Psi(u)_3$	$\Psi(u)_4$	$\Psi(u)_5$	$\Psi(u)_6$
$-\infty$	$-\infty$	−1.0000	0.000	0.00	0.00	0.00
+1.5	−1.4382	−0.9580	+0.224	+0.12	−0.17	−0.47
+2.5	−1.1203	−0.8869	+0.243	−0.09	−0.46	−0.34
+3.5	−0.8024	−0.7435	+0.085	−0.41	−0.45	+0.45
+4.5	−0.4845	−0.5068	−0.237	−0.55	+0.09	+1.14
+5.5	−0.1666	−0.1863	−0.518	−0.27	+0.73	+0.66
+6.5	+0.1513	+0.1694	−0.526	+0.25	+0.75	−0.61
+7.5	+0.4692	+0.4930	−0.253	+0.54	+0.12	−1.15
+8.5	+0.7871	+0.7344	+0.073	+0.42	−0.44	−0.50
+9.5	+1.1050	+0.8819	+0.240	+0.10	−0.47	+0.32
+10.5	+1.4229	+0.9558	+0.227	−0.11	−0.18	+0.48
+11.5	+1.7408	+0.9861	+0.138	−0.15	+0.05	+0.21
+12.5	+2.0587	+0.9964	+0.061	−0.09	+0.10	−0.02
+13.5	+2.3766	+0.9992	+0.020	−0.04	+0.06	−0.07
$+\infty$	$+\infty$	+1.0000	0.000	0.00	0.00	0.00

Tabelle XIV B.

x	$m\mathfrak{U}$	B	$\frac{1}{2}m\Phi$	$(\mathrm{B}-\mathrm{R})_0$	$F_3\Psi_3$	$(\mathrm{B}-\mathrm{R})_3$
$-\infty$	-500	-500	-500.0	0.0	0.0	0.0
$+1.5$	$+6$	-494	-479.0	-15.0	-9.4	-5.6
$+2.5$	$+36$	-458	-443.4	-14.6	-10.2	-4.4
$+3.5$	$+78$	-380	-371.7	-8.3	-3.6	-4.7
$+4.5$	$+149$	-231	-253.4	$+22.4$	$+9.9$	$+12.5$
$+5.5$	$+161$	-70	-93.1	$+23.1$	$+21.7$	$+1.4$
$+6.5$	$+183$	$+113$	$+84.7$	$+28.3$	$+22.0$	$+6.3$
$+7.5$	$+134$	$+247$	$+246.5$	$+0.5$	$+10.6$	-10.1
$+8.5$	$+114$	$+361$	$+367.2$	-6.2	-3.1	-3.1
$+9.5$	$+74$	$+435$	$+440.9$	-5.9	-10.1	$+4.2$
$+10.5$	$+34$	$+469$	$+477.9$	-8.9	-9.5	$+0.6$
$+11.5$	$+19$	$+488$	$+493.1$	-5.1	-5.8	$+0.7$
$+12.5$	$+10$	$+498$	$+498.2$	-0.2	-2.6	$+2.4$
$+13.5$	0	$+498$	$+499.6$	-1.6	-0.8	-0.8
$+\infty$	$+2$	$+500$	$+500.0$	0.0	0.0	0.0

Tabelle XIV C.

x	$F_4\Psi_4$	$(\mathrm{B}-\mathrm{R})_4$	$F_5\Psi_5$	$(\mathrm{B}-\mathrm{R})_5$	$F_6\Psi_6$	$(\mathrm{B}-\mathrm{R})_6$	$(\mathrm{B}-\mathrm{R})m\mathfrak{U}$
$-\infty$	0.0	0.0	0.0	0.0	0.0	**0.0**	
$+1.5$	-0.5	-5.1	-0.8	-4.3	-1.4	$\mathbf{-2.9}$	-2.9
$+2.5$	$+0.4$	-4.8	-2.2	-2.6	-1.0	$\mathbf{-1.6}$	$+1.3$
$+3.5$	$+1.6$	-6.3	-2.2	-4.1	$+1.3$	$\mathbf{-5.4}$	-3.8
$+4.5$	$+2.2$	$+10.3$	$+0.4$	$+9.9$	$+3.3$	$\mathbf{+6.6}$	$+12.0$
$+5.5$	$+1.1$	$+0.3$	$+3.5$	-3.2	$+1.9$	$\mathbf{-5.1}$	-11.7
$+6.5$	-1.0	$+7{,}3$	$+3.6$	$+3.7$	-1.8	$\mathbf{+5.5}$	$+10.6$
$+7.5$	-2.2	-7.9	$+0.6$	-8.5	-3.3	$\mathbf{-5.2}$	-10.7
$+8.5$	-1.7	-1.4	-2.1	$+0.7$	-1.4	$\mathbf{+2.1}$	$+7.3$
$+9.5$	-0.4	$+4.6$	-2.3	$+6.9$	$+0.9$	$\mathbf{+6.0}$	$+3.9$
$+10.5$	$+0.4$	$+0.2$	-0.9	$+1.1$	$+1.4$	$\mathbf{-0.3}$	-6.3
$+11.5$	$+0.6$	$+0.1$	$+0.2$	-0.1	$+0.6$	$\mathbf{-0.7}$	-0.4
$+12.5$	$+0.4$	$+2.0$	$+0.5$	$+1.5$	-0.1	$\mathbf{+1.6}$	$+2.3$
$+13.5$	$+0.2$	-1.0	$+0.3$	-1.3	-0.2	$\mathbf{-1.1}$	-2.7
$+\infty$	0.0	0.0	0.0	0.0	0.0	**0.0**	$+1.1$

enthält die Halbierungs- oder Wechselpunkte der in § 188 gegebenen Verteilungstafel, unter Hinzufügung der beiden Stellen $-\infty$ und $+\infty$. Daneben sind die zugehörigen Werte des Hilfsarguments

$$u = h(x - c) = 0.3179(x - 6.024)$$

angesetzt. Dann folgen die Werte von $\Phi(u)$, sowie von $\Psi(u)_3$ bis $\Psi(u)_6$, wobei die Stellenzahl mit Rücksicht auf die in Tabelle XII gegebenen Werte der Größen $mE:4$ oder F bemessen ist.

In dem Abschnitt XIV B sind der bessern Übersicht wegen zunächst die Werte von x aus XIV A wiederholt. In der zweiten Spalte $m\mathfrak{U}$ ist neben $x = -\infty$ der Betrag von $-\frac{1}{2}m$, d. h. hier -500, gesetzt, dann folgen die Werte der Verteilungstafel. Zu diesen Zahlen werden nun nach dem in § 184 Tabelle II gegebenen Schema die Werte der ersten Summenreihe gebildet, d. h. man setzt neben jedem Gliede der Spalte $m\mathfrak{U}$ die Summe aus diesem Gliede und den darüber stehenden Zahlen an. Die so entstehende, mit B überschriebene Summenspalte, die mit $\frac{1}{2}m$ schließen muß, enthält offenbar die beobachteten Werte der linken Seite von (36), wenn daselbst für die x die Wechselpunkte gesetzt werden. Damit sind die nötigen Vorbereitungen erledigt.

Da es von Interesse ist, den Einfluß der einzelnen Glieder der Φ-Reihe zu übersehen, so ist die weitere Rechnung in folgender Weise angelegt worden. Neben B stehen die Werte von $\frac{1}{2}m\Phi(u)$, die durch Multiplikation der vorher ausgeschriebenen $\Phi(u)$ mit $\frac{1}{2}m = 500$ erhalten werden. Dann werden die in Spalte $(\mathrm{B}-\mathrm{R})_0$ stehenden Differenzen

$$(\mathrm{B}-\mathrm{R})_0 = B - \tfrac{1}{2}m\,\Phi(u)$$

gebildet, die den Unterschied zwischen Beachtung und Rechnung angeben, wenn man von der Φ-Reihe nur das erste Glied berücksichtigt. Weiter folgt die Spalte der Produkte $F_3\Psi_3$, wo F_3 nach Tabelle XII gleich -41.9 ist. Die in der nächsten Spalte angesetzten Differenzen

$$(\mathrm{B}-\mathrm{R})_3 = (\mathrm{B}-\mathrm{R})_0 - F_3\Psi_3$$

geben die zwischen Beobachtung und Rechnung übrig bleibenden Unterschiede an, wenn man in der Φ-Reihe bis Φ_3 geht. In derselben Weise sind die in Tabelle XIV C aufgeführten Größen $(\mathrm{B}-\mathrm{R})_4$, $(\mathrm{B}-\mathrm{R})_5$ und $(\mathrm{B}-\mathrm{R})_6$ gebildet, indem man schrittweise auch die Produkte $F_4\Psi_4$, $F_5\Psi_5$ und $F_6\Psi_6$ heranzieht.

§ 199. Der Gang der Zahlen $(\mathrm{B}-\mathrm{R})_0$ lehrt, daß das einfache Exponentialgesetz, das ja durch das Anfangsglied der Φ-Reihe gegeben ist, den Verlauf der Zahlen B nur recht unvollkommen darstellt, und zwar kommt das nicht allein in den Beträgen der übrig bleibenden Abweichungen, sondern mehr noch in dem Verhalten der Vorzeichen zum Ausdruck. Die Hinzufügung des Gliedes $F_3\Phi_3$ bewirkt, wie die $(\mathrm{B}-\mathrm{R})_3$ lehren, eine wesentliche Verbesserung der Darstellung, womit die schon aus dem Betrage von E_3 zu entnehmende Anzeige einer ausgesprochenen Unsymmetrie der Verteilung bestätigt wird. Die folgenden Glieder drücken dann trotz der nicht sehr großen Beträge

von F_4, F_5 und F_6 die stärkeren Abweichungen noch weiter herunter und vermehren die Anzahl der Vorzeichenwechsel, so daß sich die berechnete Summenkurve mit mehrfach wechselnden Ausbiegungen von mäßiger Größe durch die beobachtete Kurve hindurchschlingt.

Die letzte mit $(\mathrm{B}-\mathrm{R})m\mathfrak{U}$ bezeichnete Spalte enthält die ersten Differenzen zu den Zahlen $(\mathrm{B}-\mathrm{R})_6$, wobei die Differenz jedesmal neben den Minuendus geschrieben ist. Diese Differenzen liefern offenbar die Unterschiede zwischen den beobachteten und den berechneten Werten von $m\mathfrak{U}(x)$.

Der Anschluß der mit sechs Elementen berechneten Verteilung an die beobachteten Zahlen ist als sehr befriedigend anzusehen, wie am besten erhellt, wenn man sich die beobachtete und die berechnete Kurve für $m\mathfrak{S}(x)$ gezeichnet denkt. Nimmt man hierbei zur Darstellung der Einheit in $m\mathfrak{S}(x)$ die Länge von 1 mm, so dehnt sich jede der beiden Kurven in der Höhe über einen Raum von 1000 mm aus, während die zwischen ihnen bestehenden Abweichungen unter 7 mm bleiben.

Wird die Verteilung als stetig vorausgesetzt, so hätte man die Vergleichung zwischen Beobachtung und Rechnung mit den für diesen Fall in § 197 angegebenen Elementen durchzurechnen. Ich unterlasse es jedoch, die Zahlen, die sich so ergeben, herzusetzen, weil der Gang der Widersprüche B — R im wesentlichen dasselbe Bild zeigt, wie im Falle der Unstetigkeit.

Weitere mit sechs numerischen Elementen durchgerechnete und sehr verschiedenen Gebieten entnommene Beispiele findet man in der Arbeit von *F. Werner* *„Beiträge zur Kollektivmaßlehre“* (*Wundt*, Philosophische Studien Band XIV). Man kann aus den daselbst mitgeteilten Zahlen entnehmen, daß die Glieder mit Φ_5 und Φ_6 für gewöhnlich von untergeordneter Bedeutung sind und höchstens bei ganz übermäßig unsymmetrischen Verteilungen eine Rolle spielen. Bei dieser Gelegenheit mag erwähnt werden, daß in den Figuren der genannten Abhandlung durch ein Versehen als Fehler der Verteilungsordinaten die zwischen Beobachtung und Rechnung übrigbleibenden Widersprüche von $m\mathfrak{S}(x)$, statt von $m\mathfrak{U}(x)$, angebracht worden sind.

§ 200. Die hier zur Erläuterung der Rechenvorschriften benutzte Auszählung der Endnullen im „Thesaurus“, die an sich als eine müßige Spielerei betrachtet werden könnte, ist mit Absicht als Beispiel gewählt worden. Schon bei früherer Gelegenheit wurde mehrfach darauf hingewiesen, daß der mathematische Inhalt der „Häufigkeitsrechnung“, d. h. also die Lehrsätze über die $\mathfrak{W}$-Größen und die Funktion $\mathfrak{S}(x)$, von der Voraussetzung zufälliger Ereignisse unabhängig sind, und daß infolgedessen die Anwendung jener Lehrsätze auf beobachtete Dinge durchaus nicht auf die mit den Einwirkungen des

Zufalls behafteten Vorkommnisse beschränkt ist. Die Frage, ob der Zufall mitspielt oder nicht wird erst dann von Belang, wenn es sich um die richtige Deutung der errechneten Ausdrücke und Zahlen handelt. Aus diesem Grunde war es offenbar nicht ohne Interesse, ein Beispiel heranzuziehen, bei dem von Zufall keine Rede ist. Dadurch wird natürlich nicht die Frage ausgeschlossen, ob und wie weit man derartige Kollektivreihen mit einem Urnenschema in Parallele stellen darf, und man kann sagen, daß gerade bei dem Beispiel der Endnullen die Vergleichung mit dem Urnenschema sehr nahe liegt. Zu dem Ende hat man eine Urne anzunehmen, die eine weiße Kugel (für die Endnull) und neun schwarze Kugeln (für die neun anderen Endziffern) enthält. Ferner ist, entsprechend den 60 Zeilen einer Thesaurus-Spalte, der einzelne Versuch aus 60 Zügen mit Zurücklegung der Kugel zusammenzusetzen. Für dieses Schema hat man, wenn

$$\mathfrak{W}(\text{Endnull}) = p = 0.1, \quad 1 - p = q = 0.9$$

geschrieben wird, nach den früheren Sätzen

$$\mathfrak{D}(u^x) = (up + q)^{60},$$

während andererseits

$$\mathfrak{D}(u^x) = \sum_x \mathfrak{U}(x) u^x$$

ist. Berechnet man nun daraufhin die $\mathfrak{U}(x)$ direkt nach der Binomialformel

$$\mathfrak{U}(x) = [60!\, p^x q^{60-x}] : [x!\,(60 - x)!]$$

von $x = 0$ bis $x = 14$, so erhält man das nachstehende Täfelchen, in dem B den beobachteten, R den berechneten Wert von $1000\,\mathfrak{U}(x)$ bedeutet.

$x =$	0	1	2	3	4	5
$B =$	0	6	36	78	149	161
$R =$	1.8	12.0	39.3	84.3	133.6	166.2
$x =$	6	7	8	9	10	11
$B =$	183	134	114	74	34	19
$R =$	169.3	145.1	106.8	68.6	38.9	19.6
$x =$	12	13	14	15—60		
$B =$	10	0	2	0		
$R =$	8.9	3.6	1.4	0.6		

Die Übereinstimmung ist nach diesen Zahlen immerhin genügend, da die Widersprüche $B - R$ von derselben Größenordnung sind, wie die aus der Formel (31) in § 130 folgenden Beträge der zugehörigen Streuungen.

§ 201. Wirft man zum Schlusse noch einen raschen Blick über die in den letzten Abschnitten behandelte Form der numerischen

Bearbeitung eines K.-G., so erkennt man, daß der Gang der Rechnung von der Verteilungstafel ab in vier deutlich geschiedene Abschnitte zerfällt, über die jetzt noch einige allgemeine Bemerkungen nachgetragen werden sollen.

Der erste Abschnitt umfaßt die Berechnung der mit der Konstante e und der Größe $y = x - e$ gebildeten Ausdrücke

$$\mathfrak{D}(y^n) \quad \text{oder} \quad m\mathfrak{D}(y^n),$$

wobei man sich vor Beginn der Rechnung darüber schlüssig zu machen hat, ob man mit den Produkten $\mathfrak{U}(x)y^n$ oder mit den Summenreihen arbeiten will. Für den, der einmal eine nicht zu kurze Verteilungstafel für drei oder mehr Elemente nach beiden Vorschriften durchgerechnet hat, wird allerdings die Wahl nicht schwer sein, denn der an sich lästige Umstand, daß bei der Summenmethode vielziffrige Zahlen auftreten, wird durch die Übersichtlichkeit der Rechnung und durch die ständigen Kontrollen mehr als ausgeglichen. Dazu kommt noch ein andrer Punkt: bei dem Produktverfahren unterliegt der Rechner leicht der Versuchung, die Verteilungstafel stark zusammenzuziehen, um die Arbeit zu sparen; wir haben aber gesehen, daß starke Abrundung durchaus vom Übel ist. Bei der Summenmethode hingegen macht es nur einen geringen Unterschied, ob die Verteilungstafel 15 oder 30 oder noch mehr notierte Argumente enthält. Die Bildung der Summenreihen gestaltet sich fast zu einer bloßen Schreibarbeit, sobald man der unter den Berufsrechnern verbreiteten Gewohnheit folgt, bei der Addition oder Subtraktion zweier Zahlen immer von links nach rechts zu rechnen.

Die besondere Anordnung des Rechenschemas, die bei dem oben behandelten Beispiel beschrieben wurde, ist aus der Absicht entsprungen, den Gebrauch der Rechenmaschine oder besonderer Hilfstafeln zu umgehen. Zu dem Ende sind außer bloßen Additionen nur Multiplikationen mit kleinen Faktoren eingeführt worden. Will man dagegen die Rechenmaschine heranziehen, was an sich durchaus nicht unzweckmäßig ist, so lassen sich noch andere brauchbare Anordnungen aufstellen, auf die jedoch hier nicht näher eingegangen werden soll.

Der zweite Abschnitt umfaßt den Übergang auf die numerischen Elemente. Daß die Potenzmittel $\mathfrak{D}(y^n)$ für sich allein zur Charakteristik eines K.-G. dienen können, ist schon früher erwähnt worden, ebenso aber auch, daß sie für diesen Zweck weniger geeignet sind, als gewisse aus ihnen gebildete Verbindungen, wie sie hier durch die D-Koeffizienten gegeben sind. Diese Koeffizienten, zusammen mit geeigneten Tafeln der Funktionen Φ_p, geben sofort ein deutliches Bild, wie das einzelne Element auf den Verlauf des K.-G. einwirkt. Bei dieser Gelegenheit mag auch noch der Umstand hervorgehoben werden, daß die Größen D_p, der linearen Transformation des Arguments

gegenüber, invariant sind. Man wird demnach, sobald man über das zweite Element str(x) hinausgehen will, stets darauf angewiesen sein, die Rechnung bis zu den nächstfolgenden D_p — oder einem gleich tauglichen System charakteristischer Bestimmungsstücke — auszudehnen.

Die hier benutzte Anordnung zur Ableitung der Elemente ist auf den Gebrauch der Logarithmentafel berechnet, die meines Erachtens in diesem Teile der Arbeit den Vorzug vor anderen Hilfsmitteln verdient. Denn man reicht für gewöhnlich mit einer vierstelligen Tafel vollkommen aus, sobald man nur bei der zweiten Form der Summenmethode die Vorsicht gebraucht, die beiden in § 191 mit e und f bezeichneten Größen möglichst nahe an $\mathfrak{D}(x)$ zu legen.

Der dritte Abschnitt der numerischen Bearbeitung beschäftigt sich mit der Ermittelung der Widersprüche, die das errechnete Elementensystem zwischen Beobachtung und Rechnung übrig läßt. Ob die Durchführung dieses Teils der Rechnung im einzelnen Falle angezeigt ist oder nicht, das hängt offenbar jedesmal von den besonderen Umständen dieses Falles ab. Wenn z. B. die beobachtete Verteilungskurve sehr stark springt, so wird man auf die Ermittelung der Widersprüche keinen besonderen Wert legen, da man im voraus auf einen unbefriedigenden Anschluß gefaßt sein muß. Andrerseits wird man, wenn eine größere Anzahl gleichartiger und regelmäßig laufender Verteilungen vorliegt, sich damit begnügen dürfen, die eine oder die andere Kurve als Stichprobe durchzurechnen.

Der vierte Abschnitt endlich, nämlich die Vergleichung mit einer theoretischen Verteilung, kommt offenbar nur bei solchen Kollektivreihen in Frage, bei denen man Grund hat, die Existenz eines bestimmten Verteilungsgesetzes, wenn auch nur hypothetisch, anzunehmen.

Bei dieser Gelegenheit mag noch die Bemerkung eingeschaltet werden, daß sich eine Verteilung, welche die $n + 1$ beobachteten Argumente $x = 0, 1, \cdots n$ enthält, jederzeit auf die Gleichungen des *Poisson*schen Schemas reduzieren läßt. Setzt man nämlich in der für eine solche Verteilung geltenden Gleichung

$$\mathfrak{D}(u^x) = \mathfrak{U}(0) + \mathfrak{U}(1)u + \cdots + \mathfrak{U}(n)u^n$$

$1 + v$ für u und entwickelt rechts nach v, so erhält man eine Gleichung von der Gestalt

$$\mathfrak{D}[(1 + v)^x] = b(0) + b(1)v + \cdots + b(n)v^n,$$

in der $b(0) = 1$ ist, da für $v = 0$ die linke Seite den Wert Eins annimmt. Zerlegt man nun das Polynom auf der rechten Seite in seine Linearfaktoren, so wird

$$\mathfrak{D}(u^x) = (1 + p_1 v)(1 + p_2 v) \cdots (1 + p_n v),$$

wo die Größen $-1 : p_k$ offenbar die Wurzeln des Polynoms bedeuten. Schreibt man weiter

$$q_k = 1 - p_k,$$

so wird

$$\mathfrak{D}(u^x) = (p_1 u + q)(p_2 u + q) \cdots (p_n u + q).$$

Das ist aber die in § 147 (10) angesetzte Ausgangsgleichung für das *Poisson*sche Schema. Selbstverständlich hat eine solche Reduktion nur dann eine mehr als bloß formale Bedeutung, wenn die Größen p_k reelle positive echte Brüche sind.

§ 202. Zusammenfassend kann man sagen, daß der hier besprochene Rechnungsgang weder besonders schwierig, noch besonders langwierig ist, immerhin aber einen gewissen Rechnungsaufwand verlangt, der dann ins Gewicht fällt, wenn Kollektivreihen in größerer Menge gleichzeitig zu untersuchen sind. Aus diesem Grunde habe ich Gewicht darauf gelegt, die Rechenvorschriften so zu gestalten, daß ein erheblicher Teil der Arbeit von Hilfskräften erledigt werden kann, die keine besondere wissenschaftliche Vorbildung besitzen. Im übrigen aber muß, da es sich doch nun einmal um die numerische Bearbeitung von Zählungen handelt, so oder so ein bestimmtes Quantum von Ziffern „vertilgt" werden, und es macht manchmal einen absonderlichen Eindruck, wenn der Bearbeiter guter Beobachtungsreihen über den erschrecklichen Ziffernverbrauch jammert und eine unsachgemäße Behandlung der Zahlen damit zu rechtfertigen sucht, daß die mathematisch korrekte Behandlung zuviel Mühe mache. Wie eine solche Ausrede zu beurteilen ist, lehrt die Vergleichung mit anderen Gebieten. Wenn z. B. jemand die mikroskopische Struktur organischer Gebilde selbständig untersuchen will, so gilt es als allererste Forderung, daß er sich mit der dazu nötigen Technik vertraut mache. Genau dasselbe gilt aber auch von der numerischen Bearbeitung beobachteter Zahlen, *denn das wissenschaftliche Rechnen ist eine Technik, die gegenwärtig in ihrer Art ebenso mannigfaltig und durchgebildet ist, wie etwa die Technik des Mikroskopierens oder der chemischen Analyse.* Daß man mit den Ziffern nicht wüsten soll, ist ja selbstverständlich, und es ist auch keine Frage, daß auf den Gebieten, auf denen die Kollektivmaßlehre als Werkzeug der Untersuchung gegenwärtig hauptsächlich in Betracht kommt, also z. B. in der Morphologie organischer Gebilde, in der Psychophysik und in der Meteorologie, vorläufig die mehr summarische Beschreibung zahlreicher gleichartiger Kollektivreihen wichtiger ist, als die penible Untersuchung eines einzelnen K.-G. Durch eine solche Einschränkung des Zieles der Rechnung reduziert sich der Rechnungsaufwand sehr erheblich. Geht man z. B. mit den numerischen Elementen nur bis D_3 oder höchstens D_4, so schrumpft der Ziffernverbrauch, wie ein Blick auf das oben behandelte

Beispiel lehrt, stark zusammen. Trotzdem bleibt der Satz bestehen, daß zur sachgemäßen Handhabung der Kollektivmaßlehre doch noch etwas mehr nötig ist, als der primitive Gebrauch der vier Spezies.

Dreiundzwanzigste Vorlesung.

Numerische Bearbeitung: Beobachtungsgleichungen.

§ 203. Die direkte Mittelbildung und die daraus unter gewissen Voraussetzungen hergeleitete Summenmethode kann als der natürlichste Rechnungsgang angesehen werden, wenn es sich um einen unstetigen K.-G. handelt. Ist dagegen der vorgelegte K.-G. stetig oder wird er im Falle der Unstetigkeit gleichwohl als stetig behandelt, so sind jene Rechenvorschriften an die Bedingung gebunden, daß die Verteilungstafel äquidistante Wechselpunkte und eine hinreichend kleine Abrundung besitze, denn andernfalls ist das Ergebnis unsicher oder auch ganz illusorisch. Nun kommt es aber öfters vor, daß der Beobachter aus Unkenntnis der für eine sachgemäße numerische Bearbeitung notwendigen Voraussetzungen gegen jene Bedingung verstößt, sei es bei der Mitteilung seiner Zählungen, sei es — was noch ungünstiger ist — schon bei der Ausführung seiner Beobachtungen. Ebenso kann es vorkommen, daß die Beschaffenheit des untersuchten K.-G. die Erfüllung der genannten Bedingung von vornherein ausschließt: Beispiele hierfür trifft man in der Psychophysik bei der sogenannten Methode der richtigen und falschen Fälle. Unter solchen Umständen wird der Rechner vor die Frage gestellt, welches Verfahren er statt des ihm nunmehr verschlossenen Weges der direkten Mittelbildung zu benutzen habe. Die Antwort hierauf beruht auf den nachfolgenden Überlegungen.

Die Beobachtung liefert für eine Reihe von Werten des Arguments x die zugehörigen Werte von $\mathfrak{S}(x)$. Man kann also, wenn man die Parameter c, h vorläufig nach Gutdünken wählt, mit dem Hilfsargument

$$u = h(x - c)$$

für die einzelnen notierten x zunächst die Gleichungen

$$2\mathfrak{S}(x) - 1 - \Phi(u) = D_1\Phi(u)_1 + D_2\Phi(u)_2 + \cdots \qquad (1)$$

ansetzen. Dieses Gleichungssystem ist nun noch zu vervollständigen, da sich seine Aussagen auf die Punkte beschränken, die dem Gebiete der möglichen Argumentwerte angehören. Zu dem Ende denken wir uns, wenn k das kleinste, g das größte mögliche Argument bedeuten,

und a eine beliebig kleine angebbare positive Größe bezeichnet, zu (1) noch die Gleichungen hinzugefügt, die den Argumentwerten

$$x = k - a,\ k - 2a,\ k - 3a,\ \ldots,$$

$$x = g + a,\ g + 2a,\ g + 3a,\ \ldots$$

entsprechen. Das so entstehende System, nach der Reihenfolge der x geordnet, nennen wir die *Summengleichungen.*

Die Summengleichungen sind, weil man von ihnen ausgehend rückwärts wieder zu der Summentafel gelangen kann, als mit dieser Tafel äquivalent anzusehen. Dieselbe Eigenschaft kommt ferner auch allen anderen Gleichungssystemen zu, die aus den Summengleichungen in der Weise hergeleitet sind, daß das neue System dem ursprünglichen jedesmal äquivalent wird. Ein Beispiel hierfür erhält man, wenn man jede Summengleichung von der nächstfolgenden abzieht. Dieses besondere System kann man als die *Verteilungsgleichungen* bezeichnen, da ja die Subtraktion jedes $\mathfrak{S}(x)$ von dem nächstfolgenden auf die Verteilungszahlen führt. Im übrigen wollen wir alle derartigen Systeme unter dem Namen *Beobachtungsgleichungen* zusammenfassen.

Das System der Summengleichungen, bei dem wir einstweilen stehen bleiben, liefert unendlich viele Bedingungen, in denen unendlich viele Unbekannte, nämlich die Größen D_p auftreten, während alles übrige bekannt ist. Mit einem solchen System ist nun einstweilen nichts rechtes anzufangen. Indessen würde auch dann, wenn ein sicheres Verfahren zur Auflösung der unendlich vielen Gleichungen angegeben werden könnte, sich die Aufgabe als unbestimmt erweisen, da ja durch die im Gebiete der möglichen x gegebenen Summenpunkte unendlich viele Summenkurven hindurchgelegt werden können, die sämtlich den allgemeinen, für diese Kurven geltenden, Grenz- und Stetigkeitsbedingungen genügen. Jede dieser Summenkurven würde aber auf ein ganz bestimmtes Wertsystem der D_p führen, so daß in Wahrheit das vorliegende Gleichungssystem unendlich viele *Lösungen* zuläßt. Es handelt sich also vor allem darum, die tatsächlich vorhandene Unbestimmtheit irgendwie aufzuheben.

§ 204. Der gewöhnlich betretene und meistens auch allein gangbare Weg zur Aufhebung der Unbestimmtheit besteht darin, daß man die D_p hinter einem gewissen Gliede, z. B. von D_5 ab, sämtlich als unmerklich annimmt und demgemäß einfach gleich Null setzt: es sind dann nur noch die vorhergehenden D_p zu ermitteln. Geht man nun die Gleichungen, die zu den außermöglichen x gehören, der Reihe nach in der Richtung nach $-\infty$ hin durch, so ist hierbei beständig $\mathfrak{S}(x) = 0$, ferner geht $\Phi(u)$ gegen -1, d. h. die linke Seite der Gleichungen geht gegen Null. Das Gleiche gilt für die rechte Seite,

da ja die $\Phi(u)_p$ in den beibehaltenen Gliedern mit unbeschränkt abnehmenden u gegen Null gehen. Die Gleichungen kommen also, je weiter man fortschreitet, der Identität $0 = 0$ immer näher. Das gleiche Ergebnis stellt sich ein, wenn man die Gleichungen in der Richtung nach $+\infty$ hin durchwandert. Da nun die numerische Rechnung immer nur mit einer beschränkten Stellenzahl arbeitet, so ist von der Gesamtheit der Summengleichungen nur eine begrenzte Anzahl für den Rechner wirklich brauchbar, nämlich die Gleichungen für die möglichen x und dazu allenfalls noch eine oder einige der beiderseits anliegenden Gleichungen. Damit ist man bei einem endlichen System von — sagen wir — n Gleichungen mit p Unbekannten angelangt. Hierbei sind nun die drei Fälle $n < p$, $n = p$ und $n > p$ zu unterscheiden.

Im ersten Falle ist die Unbestimmtheit offenbar nur vermindert, aber nicht vollständig aufgehoben; dieser Fall möge vorläufig beiseite bleiben. Im zweiten Falle ist die Aufgabe bestimmt und vorläufig erledigt, sobald man die Gleichungen nach den Unbekannten aufgelöst hat, was ja wegen der linearen Beschaffenheit der Gleichungen ohne weiteres ausführbar ist. Im dritten Falle ist die Aufgabe zunächst überbestimmt, und es wird für gewöhnlich kein Wertsystem der Unbekannten existieren, das die als brauchbar beibehaltenen Gleichungen gleichzeitig befriedigt. Oder anders ausgedrückt: jedes Wertsystem der Unbekannten, das man als *Lösung* in die Gleichungen einsetzt, wird gewisse *Widersprüche* zwischen der Beobachtung und der Rechnung übrig lassen. Damit ist der Rechner vor die neue Aufgabe gestellt, aus der Gesamtheit der Lösungen diejenige herauszusuchen, welche als die *annehmbarste* gelten soll und demgemäß bei der weiteren Benutzung der errechneten Zahlen zugrunde zu legen ist.

Die hier neu hinzutretende Aufgabe kehrt in allen messenden Wissenschaften ständig wieder und hat darum zur Ausbildung eines besonderen Algorithmus geführt, der als *Ausgleichungsrechnung* oder auch als *Interpolation* (im weiteren Sinne dieses Wortes) bezeichnet wird und sich gegenwärtig zu einem selbständigen Kapitel der angewandten Mathematik entwickelt hat. Da dieser Algorithmus, abgesehen von zahlreichen monographischen Darstellungen, jetzt fast in jedem Lehrbuche der Astronomie, Geodäsie und messenden Physik abgehandelt wird, und da überdies eine Begründung der Interpolationstheorie außerhalb des Rahmens unserer Aufgabe liegt[1]), so dürfen wir uns mit dem Satze begnügen, daß die Ausgleichungsrechnung bestimmte Wege zur Aufsuchung der annehmbarsten Lösung bietet.

§ 205. Das Rechnen mit den Beobachtungsgleichungen bietet,

1) Eine kurze Darstellung der allgemeinen Grundlagen der Interpolation habe ich in meinen „*Grundlinien des wissenschaftlichen Rechnens*“ (Leipzig 1903) gegeben.

gegen die direkte Mittelbildung und die Summenmethode gehalten, bestimmte Vorteile und ebenso bestimmte Nachteile. Ein Vorteil ist es unzweifelhaft, daß die $\mathfrak{S}(x)$ ohne weiteres so benutzt werden können, wie sie beobachtet worden sind, und daß alle Fragen, die mit der Größe der Abrundung zusammenhängen, von vornherein abgeschnitten werden. Ferner ist es ein Vorteil, daß das Verfahren unter allen Umständen anwendbar bleibt, also auch da, wo die direkte Mittelbildung versagt. Endlich ist noch die dem Rechner gewährte Bewegungsfreiheit hervorzuheben. Wenn z. B. in der Verteilungstafel für eine Teilstrecke die Beobachtung wegen irgend eines Grundes ausgefallen ist oder etwa als verdächtig ausgeschlossen werden muß, so wird dadurch die Aufsuchung der Unbekannten in keiner Weise gestört, denn der Rechner hat dann nur nötig, statt der Summengleichungen die Verteilungsgleichungen heranzuziehen, deren linke Seiten die Verteilungszahlen der einzelnen Teilstrecken enthalten, und es genügt, die Gleichung, die zu der auszuschließanden Teilstrecke gehört, einfach fortzulassen und mit den übrigen Verteilungsgleichungen weiter zu rechnen.

Als Nachteile werden gewöhnlich angesehen: erstens die Unbestimmtheit oder Willkür, die streng genommen *jeder* Ausgleichung anhaftet, zweitens die indirekte Rechnung, die aus der Auflösung der Gleichungen entspringt und die Unbekannten nicht unabhängig voneinander finden läßt. Diese Dinge sind indessen bei guten Beobachtungen von geringerer Bedeutung: wichtiger ist ein anderer Umstand. Bei der Ansetzung der Beobachtungsgleichungen ist man nämlich genötigt, die Parameter c und h, falls ihre normalen Werte nicht schon anderweitig bekannt sind, zunächst auf gut Glück anzusetzen und dann von den gefundenen D-Werten auf die Normalform überzugehen. Erfahren hierbei die ursprünglichen c, h stärkere Änderungen, so ist es angezeigt, um die vernachlässigten Glieder der Φ-Reihe möglichst unschädlich zu machen, den ganzen Ansatz nebst Ausgleichung mit den verbesserten Werten von c und h zu wiederholen; man darf also nicht immer erwarten, daß sich die gesuchte Lösung sogleich auf den ersten Wurf ergebe. Auch ist bei sperrigem Beobachtungsmaterial der Fall möglich, daß man zunächst eine imaginäre Lösung findet. Verbindet man nämlich die Gleichungen

$$D_1 = \mathfrak{D}(hc - hx), \qquad 4D_2 = 2\mathfrak{D}[(hc - hx)^2] - 1$$

mit der für jede lineare Verbindung von x geltenden Gleichung

$$\operatorname{str}(hx - hc)^2 = \mathfrak{D}[(hx - hc)^2] - \mathfrak{D}(hx - hc)^2,$$

so wird, weil

$$\operatorname{str}(hx - hc)^2 = \operatorname{str}(hx)^2 = h^2 \operatorname{str}(x)^2$$

ist,

$$h^2 \operatorname{str}(x)^2 = \frac{1}{2} + 2D_2 - D_1^2.$$

Es können also unter Umständen negative Werte von $\mathrm{str}\,(x)^2$ auftreten. Derartige Fälle bedürfen selbstverständlich einer besonderen Behandlung, für die sich keine allgemeinen Regeln aufstellen lassen.

Es soll nun noch der vorhin beiseite gelassene Fall, der eigentlich nur bei übertriebener Abrundung eintreten kann, durch einige Beispiele erläutert werden. Die Zahlen sind, wie man sehen wird, fingiert; ich bemerke jedoch ausdrücklich, daß Verteilungstafeln mit einer derartig übertriebenen Abrundung in der Literatur tatsächlich vorkommen, ohne daß ihre Urheber ein Arg dabei gehabt haben.

§ **206.** Wir betrachten zunächst eine Verteilung, die dem einfachen E.-G. folgt und auf der Gleichung

$$2\mathfrak{S}(x) - 1 = \Phi(x) \tag{2}$$

beruht. Da auf der rechten Seite die Glieder mit Φ_1 und Φ_2 fehlen, so liegt zugleich die Normalform vor und man hat offenbar

$$\mathfrak{D}(x) = 0, \quad h = 1, \quad \mathrm{str}\,(x)^2 = 0.5.$$

Dazu bilden wir nunmehr eine Verteilungstafel mit den Wechselpunkten ± 1, ± 3, ... und setzen demgemäß an:

Wechselpunkte x	-3	-1	$+1$	$+3$
$\Phi(x) = 2\mathfrak{S}(x) - 1$	-1.000	-0.843	$+0.843$	$+1.000$
$\mathfrak{S}(x)$	0.000	0.079	0.921	1.000

Die Zahlen der letzten Zeile lassen erkennen, daß außerhalb der beiden Wechselpunkte ± 3 keine vollen Teilstrecken mehr vorkommen. Damit gelangen wir zu einer Verteilungstafel, die nur drei volle Argumente enthält, nämlich

Argumente x	-2	0	$+2$
$\mathfrak{U}(x)$	0.079	0.842	0.079.

Berechnet man hieraus nach dem Verfahren der direkten Mittelbildung die Größen $\mathfrak{D}(x)$ und $\mathfrak{D}(x^2)$, so wird

$$\mathfrak{D}(x) = 0, \quad \mathfrak{D}(x^2) = 0.632, \quad \mathrm{str}\,(x)^2 = 0.632.$$

Das gefundene $\mathrm{str}\,(x)^2$ bedarf also, um den richtigen Wert 0.5 zu liefern, wegen Teilstreckenlänge und Phase der Korrektion -0.132. Nun ist die Korrektion wegen Teilstreckenlänge gleich $T^2 : 3$, also im vorliegenden Falle wegen $T = 1$ gleich -0.333, so daß der mit dem Einfluß der Phase behaftete Wert gleich $0.632 - 0.333$ oder gleich 0.299 wird. Die wegen der Phase erforderliche Korrektion ist also gleich $0.500 - 0.299 = +0.201$.

In dem vorstehenden Beispiel war wenigstens $\mathfrak{D}(x)$ richtig gefunden worden. Wir wollen nun das Beispiel abändern, indem wir jetzt

$$2\mathfrak{S}(x) - 1 = \Phi(x - 0.5)$$

setzen, wo $\mathfrak{D}(x) = 0.5$, $h = 1$ und $\operatorname{str}(x)^2 = 0.5$ wird. Mit den vier vorhin benutzten Wechselpunkten erhalten wir nunmehr die Zahlen

Wechselpunkte x	-3	-1	$+1$	$+3$
$\Phi(x) = 2\mathfrak{S}(x) - 1$	-1.000	-0.966	$+0.520$	$+1.000$
$\mathfrak{S}(x)$	0.000	0.017	0.760	1.000

aus denen die Verteilungstafel

Argumente x	-2	0	$+2$
$\mathfrak{U}(x)$	0.017	0.743	0.240

entspringt. Daraus ergibt sich wieder durch direkte Mittelbildung

$$\mathfrak{D}(x) = +0.223, \quad \mathfrak{D}(x^2) = 1.028, \quad \operatorname{str}(x)^2 = 0.978,$$
$$\operatorname{str}(x)^2 - 0.333 = 0.634.$$

Hieraus ergeben sich für $\mathfrak{D}(x)$ und $\operatorname{str}(x)^2$ die Phasenkorrektionen $+0.277$ und -0.134

Als drittes Beispiel betrachten wir eine Verteilung, bei der

$$\begin{gathered} \mathfrak{D}(x) = 0, \quad h = 1, \quad \operatorname{str}(x)^2 = 0.5, \\ D_1 = 0, \quad D_2 = 0, \quad D_3 = 0.075 \end{gathered} \tag{3}$$

ist, während die übrigen Elemente so klein sein sollen, daß sie für die dritte Dezimale in $\mathfrak{S}(x)$ nicht in Betracht kommen. Aus

$$2\mathfrak{S}(x) - 1 = \Phi(x) + 0.075\,\Phi(x)_3 = \Phi(x) + 0.3\,\Psi(x)_3 \tag{4}$$

berechnen wir ähnlich wie vorhin das Täfelchen

Wechselpunkte	-3	-1	$+1$	$+3$
$2\mathfrak{S}(x) - 1$	-1.000	-0.780	$+0.905$	$+1.000$
$\mathfrak{S}(x)$	0.000	0.110	0.952	1.000

Argumente x	-2	0	$+2$
$\mathfrak{U}(x)$	0.110	0.842	0.048

Die Verteilung ist merklich unsymmetrisch, denn aus (4) erhält man

$$\mathfrak{S}(0) = 0.42, \quad 1 - \mathfrak{S}(0) = 0.58,$$

so daß sich die Argumentmengen, die unterhalb und oberhalb des Durchschnittes $\mathfrak{D}(x) = 0$ liegen, etwa wie 2 zu 3 verhalten. Die Rechnung mit direkter Mittelbildung liefert

$$\mathfrak{D}(x) = -0.124, \quad \mathfrak{D}(x^2) = 0.632, \quad \operatorname{str}(x)^2 = 0.617,$$
$$\operatorname{str}(x)^2 - 0.333 = 0.284,$$

woraus wiederum ein sehr beträchtlicher Einfluß der Phase folgt.

§ 207. Wir wollen nun einmal annehmen, daß die $\mathfrak{S}(x)$ der drei beobachteten Beispiele ohne Angabe ihrer Herkunft als Beobachtungen eines stetigen K.-G. vorgelegt worden seien. Da ferner die Verteilungstafeln jedesmal nur drei volle Argumente enthalten, so

wollen wir versuchen, ob man nicht schon mit den beiden niedrigsten numerischen Elementen und dem einfachen E.-G. bei der Darstellung der gegebenen Zahlen auskommen könne. Bei der Lösung der gestellten Aufgabe ist man, weil die starke Abrundung den Gebrauch der direkten Mittelbildung illusorisch macht, auf die für die Wechselpunkte anzusetzenden Beobachtungsgleichungen

$$2\mathfrak{S}(x) - 1 = \Phi[h(x - c)] \tag{5}$$

angewiesen, deren linke Seiten alles enthalten, was sich auf Grund der gegebenen Zahlen über die vorgelegten Verteilungen aussagen läßt.

Für die beiden ersten Beispiele ist es natürlich selbstverständlich, daß die Benutzung der Gleichungen (5) wieder zu den als Ausgangspunkt genommenen Elementen führen wird. So liefert z. B. der zweite Fall aus den beiden Gleichungen

$$- 0.966 = 2\mathfrak{S}(-1) - 1 = \Phi[h(-1 - c)],$$
$$+ 0.520 = 2\mathfrak{S}(+1) - 1 = \Phi[h(+1 - c)]$$

mit Hilfe einer fünfstelligen Φ-Tafel zu den gegebenen Funktionswerten die Argumente

$$h(-1 - c) = -1.4991, \quad h(+1 - c) = +0.4994.$$

Daraus folgt

$$2h = 1.9985, \quad 2hc = +0.9997, \quad h = 0.9992, \quad c = +0.5000,$$

wo die kleine Abweichung in h aus den Abrundungsfehlern in den berechneten $\mathfrak{S}(x)$ entstanden ist.

Das dritte Beispiel in derselben Weise behandelt liefert mit den beiden mittleren $\mathfrak{S}(x)$ die Gleichungen

$$- 0.780 = \Phi[h(-1 - c)], \quad + 0.905 = \Phi[h(+1 - c)],$$
$$h(-1 - c) = -0.8673, \quad h(+1 - c) = +1.1806,$$
$$2h = 2.0479, \quad 2hc = -0.3133, \quad h = 1.024, \quad c = -0.153.$$

Berechnet man damit noch die Größen $2\mathfrak{S} - 1 = \Phi$ für die beiden äußeren Wechselpunkte, so wird

$$h(-3 - c) = -2.9152, \quad h(+3 - c) = +3.2285,$$
$$2\mathfrak{S}(-3) - 1 = -1.000, \quad 2\mathfrak{S}(+3) - 1 = +1.000.$$

Hiernach schließt sich das einfache E.-G. den vier gegebenen Summenpunkten innerhalb der angewandten Rechnungsschärfe vollständig an, obgleich diese Punkte einer ausgesprochen unsymmetrischen Verteilung entnommen waren. Dieser Umstand ist viel wesentlicher, als der ziemlich große Fehler von $\mathfrak{D}(x)$, denn man erkennt daraus, welches Maß von Unbestimmtheit in dem Verlaufe der Summenkurve durch die starke Abrundung der Tafelargumente erzeugt wird.

Die Beispiele sind nach Ausweis der vorhandenen Literatur, wie schon bemerkt wurde, keineswegs besonders kraß gewählt. Auch darf man nicht vergessen, daß die niedrigsten Elemente, die hier allein berücksichtigt wurden, gegen die Phase weniger empfindlich sind, als die Elemente mit höherer Nummer.

§ 208. Wenn nach dem Gesagten die starke Abrundung des Arguments vom Übel ist, so kann man fragen, welche Umstände denn eigentlich dazu nötigen, Verteilungstafeln mit unzulässig starker Abrundung aufzustellen und zu benutzen. Hierbei sind zwei Fälle zu unterscheiden, je nachdem in der Urliste die Messungen des veränderlichen Merkmals mit größerer oder geringerer Schärfe notiert worden sind.

Um die Vorstellung zu fixieren, wollen wir annehmen, daß das Argument eine Länge bedeute, und daß die gemessenen x in der Urliste mit vollen Millimetern notiert worden seien; ferner sei der Umfang der Urliste gleich 200 und der Spielraum, innerhalb dessen sich die beobachteten x bewegen, gleich 100 mm. Dann kommen auf die einzelne Teilstrecke von 1 mm Ausdehnung im Durchschnitt zwei Glieder, ferner hat man, wenn die Verteilung ungefähr dem Exponentialgesetz folgt, in der „dichtesten“ Teilstrecke etwa zehn Glieder zu erwarten. Stellt man unter solchen Umständen die Verteilungstafel her und konstruiert wie früher angegeben den Linienzug der Verteilungspunkte, so werden die Zahlen der Tafel und die Verteilungspunkte stark springen, dagegen wird der „Gang“ der Verteilung mehr oder minder deutlich zum Vorschein kommen, wenn man z. B. je zehn Teilstrecken zusammenlegt, also die x in vollen Zentimetern notiert. Gegen das Motiv einer solchen Zusammenlegung, nämlich Sichtbarmachen des Ganges, ist nun nichts einzuwenden, so lange es sich nur um eine vorläufige Orientierung handelt, dagegen wird es hinfällig, wenn man an die Berechnung der Elemente geht. Denn der Gang kommt auch bei den Millimeter-Teilstrecken vollkommen deutlich zum Vorschein, wenn man statt der Verteilungstafel die Summentafel und den Linienzug der Summenpunkte konstruiert; wichtiger ist aber noch, daß — um bei den angenommenen Zahlen zu bleiben — die Zusammenlegung in Zentimeterstrecken nicht weniger als neun Zehntel der tatsächlich beobachteten Summenpunkte beseitigt und einfach als nichtexistierend behandelt. Wenn also mit einer ursprünglich sehr kleinen Abrundung nachträglich eine Vergrößerung vorgenommen wird, so hat sich diese nach ganz anderen Gesichtspunkten zu richten, als nach der Rücksicht auf ein mehr oder minder deutliches Sichtbarmachen des Ganges der Verteilungszahlen.

Häufig spricht übrigens bei der Zusammenziehung der Verteilungstafeln auch der Wunsch mit, an Rechenarbeit zu sparen. In der

Tat macht es, wenn die Potenzmittel direkt als Produktsummen in der Gestalt

$$\sum m \mathfrak{U}(x)(x-e)^n$$

berechnet werden, einen merklichen Unterschied, ob man es mit 10 oder 20 oder gar 30 Verteilungszahlen zu tun hat, und man stößt in der Literatur gelegentlich auf die mit einem stillen Seufzer verbundene Bemerkung, daß schon die Berechnung der Größen $\mathfrak{D}(x)$ und str (x) bei längeren Verteilungstafeln eine recht beschwerliche Sache sei. In dieser Beziehung wurde aber schon früher bemerkt, daß bei der Summenmethode in ihrer zweiten Form die Länge der Verteilungstafel für den gesamten Arbeitsaufwand von verhältnismäßig untergeordneter Bedeutung ist.

In dem soeben betrachteten Falle ist die Abrundung in der Urliste klein, und man kann, wenn die Verteilungstafel unzulässig stark zusammengezogen worden war, den Fehler nachträglich durch ein Zurückgehen auf die Urliste wieder gut machen. Es kann nun aber auch der Fall vorliegen, daß schon bei der Beobachtung die Argumentwerte mit einer übermäßigen Abrundung notiert worden sind, daß also die daraus bei der Berechnung der Elemente erwachsende Schwierigkeit nachträglich überhaupt nicht mehr zu beseitigen ist. Wenn z. B. die x wieder mit vollen Millimetern notiert sind, während der ganze Spielraum des Arguments nur 3 mm bis 4 mm beträgt, so ist, wie wir vorhin an Zahlenbeispielen gesehen haben, mit einer solchen Kollektivreihe kaum etwas anzufangen. Manchmal ist es allerdings nicht eine übel angebrachte Bequemlichkeit des Beobachters, sondern eine auf den ersten Anschein ganz plausible Überlegung, die zu einer übermäßigen Zusammenziehung der Verteilungstafel Anlaß gibt. Es heißt z. B.: die einzelne Messung ist mit einem nicht genauer zu bestimmenden Fehler behaftet, der im Durchschnitt etliche Zehntelmillimeter beträgt und gelegentlich nahe an das volle Millimeter heranreicht, folglich ist ohne Rücksicht auf den Spielraum des Arguments nur das volle Millimeter zu notieren, weil sich nur dieses verbürgen läßt. Eine solche Überlegung ist ganz gut zulässig, solange sie nur auf eine *isolierte* Messung bezogen wird, sie ist dagegen verfehlt, wenn eine *Reihe* von Messungen vorliegt, denn sie läßt die Möglichkeit außer acht, daß die Wirkung der Messungsfehler durch geeignete Kombination der Messungen herabgedrückt werden kann. Darum gilt denn auch in den eigentlich messenden Wissenschaften seit langer Zeit und aus guten Gründen die Regel, daß in den notierten Zahlen die Einheit der letzten Stelle kleiner als der durchschnittlich zu befürchtende Messungsfehler sein soll.

Die vorstehenden Bemerkungen führen zu der Frage, welche Rolle bei stetigen Kollektivreihen die den beobachteten Argumentwerten anhaftenden Messungsfehler spielen.

§ 209. Bei der Untersuchung der gestellten Frage wollen wir die als gemessen notierten Argumentwerte mit x, die von den Messungsfehlern befreiten oder wahren Werte mit X bezeichnen. Dann ist $y = x - X$ der Messungsfehler und

$$x = X + y. \tag{6}$$

Weiter wollen wir die gewöhnlich zutreffende Voraussetzung machen, daß in der Urliste die X und y unabhängig voneinander variieren, und daß hierbei nicht nur X, sondern auch y einem gewissen Verteilungsgesetze folge. Die Existenz von „Fehlergesetzen" brauchen wir hier nicht weiter zu erörtern, indem wir uns mit der geschichtlichen Tatsache begnügen, daß die Fehlermengen zu den Kollektivreihen gehören, die am frühesten untersucht worden sind, und zwar lange bevor *Fechner* das Problem einer allgemeinen Kollektivmaßlehre zu formulieren unternahm.

Die Gleichung (6) bedeutet unter den gemachten Voraussetzungen eine Argumentmischung, so wir daß die Entwicklungen aus der XIII. Vorlesung einfach herübernehmen dürfen. Setzt man

$$\mathfrak{D}(x) = c, \quad \mathfrak{D}(X) = C, \quad \mathfrak{D}(y) = e, \tag{7}$$

wo das $\mathfrak{D}$-Zeichen sich immer auf das dahinter stehende Argument und seine Verteilung bezieht, so folgt aus (6) zunächst

$$c = C + e, \tag{8}$$

und weiter $x - c = (X - c) + (y - e)$, woraus durch Quadrieren und Mittelnehmen die Beziehung

$$\operatorname{str}(x)^2 = \operatorname{str}(X)^2 + \operatorname{str}(y)^2 \tag{9}$$

fließt. Setzt man ferner

$$1 = 2h^2 \operatorname{str}(x)^2 = 2H^2 \operatorname{str}(X)^2 = 2k^2 \operatorname{str}(y)^2,$$

so wird

$$h^{-2} = H^{-2} + k^{-2}, \tag{10}$$

woraus mit den Abkürzungen

$$G = h : H, \quad g = h : k \tag{11}$$

die Relation

$$G^2 + g^2 = 1 \tag{12}$$

entsteht. Endlich bilden wir noch

$$D(c, h)_q = \mathfrak{D}[\mathfrak{R}(hx - hc)_q], \quad D(C, H)_q = \mathfrak{D}[\mathfrak{R}(HX - HC)_q],$$
$$D(e, k)_q = \mathfrak{D}[\mathfrak{R}(ky - ke)_q],$$

$$N = \exp[-2hxw + 2hcw - w^2], \quad \mathfrak{D}(N) = \sum_q D(c, h)_q (2w)^q,$$

$$P = \exp[-2HX \cdot Gw + 2HC \cdot Gw - G^2w^2], \quad \mathfrak{D}(P) = \sum_q D(C, H)_q (2Gw)^q,$$

$$Q = \exp[-2ky \cdot gw + 2ke \cdot gw - g^2w^2], \quad \mathfrak{D}(Q) = \sum_q D(e, k)_q (2gw)^q.$$

Da $N = P \cdot Q$ ist, und da ferner X und y unabhängig voneinander variieren, so ist $\mathfrak{D}(N)$ gleich dem Produkte $\mathfrak{D}(P)\mathfrak{D}(Q)$, woraus

$$\sum_q D(c, h)_q (2w)^q = [\sum_q D(C, H)_q (2Gw)^q] \cdot [\sum_q D(e, k)_q (2gw)^q] \quad (13)$$

folgt. Multipliziert man die Reihen rechter Hand aus und spaltet darauf nach w, so entsteht die Beziehung

$$D(c, h)_q = \sum_r G^{q-r} D(C, H)_{q-r} \cdot g^r D(e, k)_r, \quad (14)$$

worin r bei der Summation von 0 bis q zu laufen hat.

Die vorstehende Gleichung (14) vereinfacht sich für die niedrigeren Werte von q erheblich, denn die D-Koeffizienten gehören sämtlich der Normalform an, so daß die D_1 und D_2 verschwinden. Schreibt man für den Augenblick statt $D(c, h)$, $D(C, H)$, $D(e, k)$ die Buchstaben K, L, M, so wird mit Fortlassung der für $q = 0, 1, 2$ entstehenden Identitäten

$$\left.\begin{array}{ll} K_3 = G^3 L_3 + g^3 M_3, & K_4 = G^4 L_4 + g^4 M_4, \\ K_5 = G^5 L_5 + g^5 M_5, & K_6 = G^6 L_6 + G^3 L_3 \cdot g^3 M_3 + g^6 M_6, \\ & \text{usw.} \end{array}\right\} \quad (15)$$

Wir wollen nun zusehen, wie sich an der Hand der vorstehenden Formeln die Untersuchung der Fälle gestaltet, die zu der ganzen Fragestellung Anlaß gegeben haben.

§ 210. Um die Vorstellung zu fixieren denken wir uns, daß der Spielraum des beobachteten Arguments x im ganzen 3 mm betrage, und daß man die Zehntelmillimeter noch ablesen oder doch schätzen könne, wenn auch mit einem Fehler, der im Durchschnitt auf mehrere Zehntelmillimeter steigt. Setzt man nun nach dem Grundsatze, daß nur verbürgte Beträge notiert werden sollen, die x mit vollen Millimetern an, so geht die Verteilungstafel nur über drei oder höchstens vier volle Argumente. Mit einer solchen Tafel ist aber, wie wir oben an Zahlenbeispielen gesehen haben, wegen der Phasenwirkung nichts ernstliches anzufangen, es sei denn, daß das analytische Bildungsgesetz von $\mathfrak{S}(x)$ bereits anderweit bekannt ist und nur von zwei oder höchstens drei Parametern abhängt. Wird dagegen das bei der Messung notierte Zehntelmillimeter beibehalten, so entsteht eine Verteilungstafel, die sich zunächst über mindestens dreißig Teilstrecken ausdehnt und bei genügendem Umfange des betrachteten K.-G. den Phaseneinfluß völlig sicher zu eliminieren gestattet. Es handelt sich also nur noch darum zuzusehen, wie weit man die zu den x gefundenen numerischen Elemente auf ihre wahren Werte reduzieren oder von dem Einfluß der y befreien könne.

Die in (8) angesetzte Gleichung $c = C + e$ lehrt, daß die Größe e gerade so wirkt, als ob alle Messungen von x den gemeinsamen

konstanten Fehler e besitzen. Einen solchen Fehler muß man, mag nun das benutzte Messungsverfahren mehr oder minder vollkommen sein, offenbar stehen lassen, so lange er nicht durch eine besondere Untersuchung anderweit ermittelt oder als belanglos nachgewiesen worden ist; er scheidet bei der Frage, mit der wir es hier zu tun haben, aus.

Um die anderen Formeln zu benutzen muß man zunächst den Betrag von $\operatorname{str}(y)$ kennen. Das ist nun meistens sehr wohl zu erreichen. Wird z. B. an einem beliebig herausgegriffenen Gliede der vorgelegten Kollektivreihe, dessen X den Wert A besitzen möge, nicht bloß eine einzige, sondern eine ganze Reihe von Messungen ausgeführt, so liefern die so entstehenden Gleichungen $x = A + y$, wenn sie für sich als ein K.-G. behandelt werden, die für diesen K.-G. geltende Beziehung $\operatorname{str}(x) = \operatorname{str}(y)$, da ja A hier die Rolle einer Konstante spielt; die Nebenmessung führt also zur Kenntnis von $\operatorname{str}(y)$. Ein anderer Weg besteht darin, daß man eine genügende Menge der Glieder des betrachteten K.-G. zweimal unabhängig voneinander mißt. Bezeichnet man die x, y der zweiten Messung mit x', y', so liefern die zu jedem einzelnen Gliede gehörigen Messungen die Gleichungen

$$x - x' = y - y'.$$

Werden diese Gleichungen für sich als eine Kollektivreihe behandelt, so entsteht nach den Mischungssätzen die Beziehung

$$\operatorname{str}(x - x')^2 = \operatorname{str}(y)^2 + \operatorname{str}(y')^2 = 2 \operatorname{str}(y)^2,$$

aus der wiederum $\operatorname{str}(y)$ folgt. Aber auch wenn man solche Nebenmessungen nicht heranziehen kann, wird selbst die Vernachlässigung der von y abhängenden Glieder immer noch Werte liefern, die den Namen einer Annäherung verdienen. Wird z. B. bei dem hier angenommenen K.-G. der Spielraum von y schätzungsweise gleich 1 mm, also gleich einem Drittel des Spielraums von x gesetzt, so erhält man, wenn für $\operatorname{str}(x)$ kurz s geschrieben wird, mit der Schätzung $\operatorname{str}(y) = s : 3$ nach den Gleichungen (9) und (11) die Werte

$$\operatorname{str}(X) = 0.94 \operatorname{str}(x), \quad g = 0.33, \quad G = 0.94.$$

Die Vernachlässigung der y fälscht also selbst in einem so ungünstig liegenden Falle die Streuung nur um 6 Prozent. Ähnlich liegt die Sache nach (14) bei den Größen $L_3, L_4, \ldots$, da die Potenzen von g kleine Brüche sind, während die Potenzen von G nahe bei Eins bleiben.

Die vorstehenden Bemerkungen lassen erkennen, daß die Größe der Messungsfehler kein ausreichender Grund ist, eine Verteilungstafel derart zusammenzuziehen, daß die Wirkung der Abrundungsphase störend wird. Im übrigen empfiehlt es sich natürlich, die Messungen eines K.-G. so einzurichten, daß der Messungsfehler auf die numerischen

Elemente überhaupt keinen merklichen Einfluß ausübt, denn die Benutzung der oben entwickelten Reduktionen bleibt immer ein Notbehelf, der überdies mit einer Vermehrung der Arbeit verbunden ist.

§ 211. Die früheren Erörterungen über die direkte Mittelbildung und die Beobachtungsgleichungen gingen immer davon aus, daß jedesmal nur einer der beiden Wege in Betracht komme. Es besteht indessen kein Hindernis, die beiden Methoden, wenn es aus irgend welchen Gründen zweckmäßig erscheint, miteinander zu kombinieren. Als Beispiel möge der Fall dienen, daß bei einem K.-G., der sich nach der Summenmethode behandeln läßt, zunächst nur die beiden niedrigsten Elemente, nämlich $\mathfrak{D}(x)$ und $\operatorname{str}(x)$, nach dieser Methode berechnet worden seien, und daß man der weiteren Rechnung den Ansatz mit den Beobachtungsgleichungen zugrunde legen wolle. Man kann dann in den Beobachtungsgleichungen für die Parameter c, h sogleich ihre normalen Werte einführen und darf daraufhin die beiden Unbekannten D_1 und D_2 von vornherein gleich Null setzen, so daß nur noch die folgenden D-Koeffizienten zu bestimmen sind. Zur besseren Erläuterung will ich die nachstehende aus 1444 Gliedern bestehende Verteilungstafel heranziehen.

Tabelle XV.

$x =$	0	1	2	3	4	5	6
$m\mathfrak{U}(x) =$	4	16	26	53	103	247	343
$x =$	7	8	9	10	11	12	13
$m\mathfrak{U}(x) =$	294	208	85	26	22	11	6

Geht man nach der ersten Form der Summenmethode bis zur dritten Summenreihe, so erhält man mit Unterdrückung der Zwischenrechnung

$$\mathfrak{D}(x) = 6.3234, \quad \operatorname{str}(x)^2 = 3.8033 \text{ (unverbessert)},$$
$$\operatorname{str}(x) = 1.929 \text{ (verbessert)}, \quad h = 0.3666.$$

Die Streuung ist wegen der Teilstreckenlänge verbessert worden, weil es sich um einen stetigen K.-G. handelt. Für das Hilfsargument erhält man mit den normalen Werten den Ausdruck

$$u = 0.3666\,(x - 6.3234),$$

worin die Wechselpunkte von $x = -0.5$ bis $x = +13.5$ zu gehen haben. Die weitere Rechnung findet man unter Fortlassung einiger für das Verständnis unwesentlicher Zwischenglieder in Tabelle XVI zusammengestellt. Unter x sind die Wechselpunkte angegeben, unter u die Werte des Hilfsarguments, dann folgen die Werte der Größen

$$B = m\mathfrak{S}(x) - \tfrac{1}{2}m, \quad \tfrac{1}{2}m\,\Phi(u),$$
$$\mathrm{I} = B - \tfrac{1}{2}m\,\Phi(u), \quad \Psi_3 = \Phi(u)_3 : 4, \quad \Psi_4 = \Phi(u)_4 : 8.$$

Tabelle XVI.

x	u	B	$\frac{1}{2}m\Phi$	I	Ψ_3	Ψ_4	II
− 0.5	− 2.5015	− 722	− 721.7	− 0.3	+ 0.0124	+ 0.0258	− 2.4
+ 0.5	− 2.1349	− 718	− 720.1	+ 2.1	+ 0.0480	+ 0.0772	− 4.3
+ 1.5	− 1.7683	− 702	− 713.0	+ 11.0	+ 0.1300	+ 0.1424	− 0.8
+ 2.5	− 1.4017	− 676	− 687.7	+ 11.7	+ 0.2317	+ 0.1031	+ 3.2
+ 3.5	− 1.0351	− 623	− 618.5	− 4.5	+ 0.2208	− 0.1714	+ 9.6
+ 4.5	− 0.6685	− 520	− 473.3	− 46.7	− 0.0383	− 0.5080	− 4.8
+ 5.5	− 0.3019	− 273	− 238.7	− 34.3	− 0.4211	− 0.4381	+ 1.9
+ 6.5	+ 0.0647	+ 70	+ 52.6	+ 17.4	− 0.5571	+ 0.1087	+ 8.4
+ 7.5	+ 0.4313	+ 364	+ 330.7	+ 33.3	− 0.2941	+ 0.5309	− 10.5
+ 8.5	+ 0.7979	+ 572	+ 534.8	+ 37.2	+ 0.0816	+ 0.4113	+ 3.3
+ 9.5	+ 1.1645	+ 657	+ 650.1	+ 6.9	+ 0.2488	+ 0.0487	+ 2.9
+ 10.5	+ 1.5311	+ 683	+ 700.1	− 17.1	+ 0.1996	− 0.1399	− 5.6
+ 11.5	+ 1.8977	+ 705	+ 716.7	− 11.7	+ 0.0955	− 0.1228	− 1.6
+ 12.5	+ 2.2643	+ 716	+ 721.0	− 5.0	+ 0.0310	− 0.0550	− 0.5
+ 13.5	+ 2.6309	+ 722	+ 721.9	+ 0.1	+ 0.0072	− 0.0158	+ 1.4

Die Bedeutung der Spalte II ist nachher anzugeben. Multipliziert man die vorläufig in der Gestalt

$$2\mathfrak{S}(x) - 1 = \Phi(u) + D_3\Phi(u)_3 + D_4\Phi(u)_4 + \cdots$$

angesetzten Summengleichungen mit $\frac{1}{2}m$ und nimmt das Glied mit $\Phi(u)$ auf die linke Seite, so nehmen die Gleichungen die Gestalt

$$\mathrm{I} = F_3\Psi_3 + F_4\Psi_4 + \cdots \tag{16}$$

an. Die Größen I sind also die Widersprüche, welche zwischen Beobachtung und Rechnung übrig bleiben, wenn man zur Darstellung von $\mathfrak{S}(x)$ nur das Anfangsglied der Φ-Reihe oder das einfache Exponentialgesetz benutzt, also die Glieder mit $\Psi_3, \ldots$ unterdrückt. Diese Widersprüche sind recht ansehnlich, müssen jedoch, wie schon ohne weitere Rechnung aus dem übereinstimmenden Gange der Zahlen I und Ψ_4 zu entnehmen ist, eine erhebliche Verkleinerung erfahren, wenn man das Glied mit Ψ_4 heranzieht.

Beschränkt man sich hier auf die beiden Glieder mit Ψ_3 und Ψ_4, so sind für die beiden Unbekannten F_3 und F_4 zwei sogenannte Normalgleichungen zu bilden. Der Kürze halber soll hier ein ganz kunstloses, aber für unseren Zweck ausreichendes, Verfahren benutzt werden. Man bringt in sämtlichen Gleichungen den Koeffizienten der Unbekannten F_3 auf positives Vorzeichen, indem man wo nötig die Gleichungen mit − 1 multipliziert; darauf summiert man die solchergestalt abgeänderten Gleichungen. Ferner wiederholt man diese Opera-

tion, indem man die Koeffizienten von F_4 auf positives Vorzeichen bringt und summiert. Dadurch entstehen die beiden Normalgleichungen

$$+ 60.7 = + 2.6172\, F_3 + 0.6101\, F_4,$$
$$+ 238.5 = - 0.1934\, F_3 + 2.8991\, F_4,$$

aus denen $F_3 = + 3.95$, $F_4 = + 82.54$ folgt. Bildet man nunmehr die in Spalte II angesetzten Größen

$$\mathrm{II} = \mathrm{I} - F_4 \Psi_4,$$

so erhält man die Widersprüche, welche übrig bleiben, wenn man bei der Darstellung von $\mathfrak{S}(x)$ außer $\Phi(u)$ auch das Glied mit Ψ_4 berücksichtigt. Die Vergleichung von I und II lehrt, daß gerade die großen Beträge von I erheblich reduziert worden sind. Ferner zeigt ein Blick auf die Zahlen Ψ_3 und II, daß die Berücksichtigung von Ψ_3 den erreichten Anschluß nicht wesentlich ändern würde.

Der hier behandelte K.-G. findet sich einer Arbeit von *H. Laurent* „*Sur la méthode des moindres carrés*" (Liouville J. Série III Tome I 1875) und ist aus einer Reihe von Beobachtungsfehlern entstanden, die bei wiederholter Messung eines Winkels vorgekommen waren. Die a. a. O. ganz ungenügend behandelte Reihe verfolgte den Zweck zu prüfen, wie weit bei Beobachtungsfehlern das einfache E.-G. geeignet sei, die Verteilung solcher Fehler darzustellen. Das starke Hervortreten des von Φ_4 abhängenden Gliedes und das Vorzeichen von F_4 läßt vermuten, daß man es in Wahrheit mit einem Gemisch von zwei oder mehr ungleichartigen Reihen zu tun hat.

§ 212. Wenn die gegebene Summentafel derart beschaffen ist, daß man zwischen den Wechselpunkten ohne Bedenken geradlinig interpolieren kann, so läßt sich dem Ansatz (1) noch eine andere Gestalt geben. Man denke sich für das Hilfsargument $u = h(x - c)$ ein bestimmtes Wertsystem ein für allemal vorgeschrieben, daraus die entsprechenden Werte von x berechnet und zu diesen die Werte von $\mathfrak{S}(x)$ aus der Summentafel interpoliert. Dann besitzen in den Gleichungen

$$m\mathfrak{S}(x) - \tfrac{1}{2}m - \tfrac{1}{2}m\,\Phi(u) = \tfrac{1}{2}m\,[D_1\Phi(u)_1 + D_2\Phi(u)_2 + \cdots] \qquad (17)$$

die Φ-Größen feststehende Werte: es fällt also das jedesmalige Ausschreiben dieser Größen aus den Tafeln fort, ferner läßt sich in den Ausdrücken, welche die Unbekannten als lineare Verbindungen der linken Seiten von (17) darstellen, das System der Koeffizienten ein für allemal aufstellen, so daß auch die Auflösung der Gleichungen erheblich abgekürzt wird. Sind überdies die vorgeschriebenen u paarweise entgegengesetzt gleich, so liefert die Summe und Differenz der zu einem u-Paare gehörigen Gleichungen jedesmal zwei neue Gleichungen, die nur die geraden oder ungeraden Unbekannten ent-

halten, so daß dadurch eine weitere Erleichterung der Auflösung eintritt.

Der vorstehend angedeutete Weg beruht auf demselben Grundgedanken, der auch bei anderen Interpolationsaufgaben, z. B. bei der Darstellung periodischer Vorgänge durch eine trigonometrische Reihe, Verwendung findet.

Vierundzwanzigste Vorlesung.

Numerische Bearbeitung: Verteilungstafeln.

§ 213. Die in den letzten Vorlesungen enthaltenen Erörterungen über die Methoden zur Bestimmung der numerischen Elemente ließen erkennen, daß der Beobachter bei der Zusammenstellung der Urliste deren spätere Bearbeitung sowohl erschweren, als auch erleichtern kann. Daraus folgt erstens, daß der Beobachter von vornherein eine deutliche Vorstellung über die zweckmäßigste Gestalt der späteren Bearbeitung haben soll, und zweitens, daß dem Rechner jede unnütze Erschwerung seiner Arbeit zu ersparen ist. Diese an sich selbstverständlichen Forderungen bleiben nun nach Ausweis der Literatur häufig unerfüllt, und zwar vorwiegend bei stetigen Kollektivreihen, während bei unstetigen Reihen verhältnismäßig selten dagegen verstoßen wird. Meistens wird — aus Unkenntnis — durch übertriebene Abrundung gefehlt; ferner trifft man bei den Verteilungstafeln mitunter auf Ansätze, die eine schiefe Auffassung von der Bedeutung der in der Tafel enthaltenen Zahlen verraten. Es wird daher nicht überflüssig sein, die Punkte, auf die es hierbei ankommt, an Zahlenbeispielen nochmals vorzuführen. Daran lassen sich dann ungezwungen die Bemerkungen über die Transformation des Arguments auf numerischem Wege und über den Zweck solcher Transformationen schließen.

Als Beispiel wähle ich die in der „*Kollektivmaßlehre*“ von *Fechner-Lipps* auf Seite 105 mitgeteilte Reihe, die sich auf die Länge des obersten Gliedes von 217 sechsgliedrigen Roggenhalmen bezieht. Die Einheit der a. a. O. angegebenen x ist 5 mm; angesetzt sind noch die Zehntel dieser Einheit. Des bequemeren Schreibens wegen wollen wir die Einheit zehnmal kleiner nehmen, was darauf hinauskommt, daß man die von *Fechner* gegebenen Zahlen mit 10 multipliziert. Man erhält dann die nachstehend in Tabelle XVII gegebene Urliste, deren Glieder bereits der Größe nach geordnet sind. Die Mehrzahl der notierten x kommt nur einmal vor; das am häufigsten beobachtete x ist 837, welches viermal auftritt. Die Länge der Teilstrecken ist jetzt 1, die dem ersten und letzten Gliede von außen anliegenden Wechselpunkte sind 428.5 und 1122.5; die Tafel dehnt sich also

über 694 Teilstrecken aus, so daß eine ziemlich starke Zusammenziehung gestattet ist, da es ja nur darauf ankommt, den Einfluß der Abrundungsphase unmerklich zu halten. Wir wollen hier je 40 Teilstrecken zusammenlegen; man erhält damit in der reduzierten Tafel 18 Teilstrecken, was bei einer Kollektivreihe wie der vorliegenden genügt.

Tabelle XVII.

429	497	528	556	576	589	590	614	619	622
623	630	641	643	655	674	677	678	681	683
689	696	699	705	714	720	720	721	725	729
737	739	741	748	748	751	751	752	756	758
758	761	762	762	764	764	767	770	772	775
776	777	779	780	781	781	784	788	790	794
800	800	804	807	809	809	813	819	820	820
821	821	823	823	823	824	828	830	831	834
837	837	837	837	839	839	846	854	855	857
858	859	860	860	862	863	868	868	869	870
870	870	871	871	874	874	875	878	879	879
880	880	883	886	888	889	889	892	892	893
893	894	897	897	899	899	900	902	902	902
904	905	906	907	907	907	912	913	914	917
919	919	920	920	923	928	930	930	931	933
934	935	935	937	944	946	946	947	957	958
958	959	960	961	962	963	965	968	969	970
971	975	975	976	977	978	979	980	982	986
988	990	990	992	993	994	995	1003	1005	1008
1009	1010	1011	1013	1015	1019	1022	1023	1027	1028
1033	1034	1040	1042	1044	1053	1055	1056	1058	1060
1062	1063	1080	1100	1112	1120	1122	—	—	—

In der Literatur begegnete man nun bei der Mitteilung von Verteilungstafeln häufig einer Vorliebe für runde Argumente. Es wird z. B., um bei dem vorliegenden K.-G. zu bleiben, die Argumentreihe 440, 480, ... aufgestellt und darnach die Menge der zu jedem Argument gehörigen Glieder aus der Urliste herausgesucht. Bildet man statt dessen sogleich die Summentafel und beachtet, daß zu der genannten Argumentreihe die Wechselpunkte 420, 460, ... gehören, so erhält man durch Abzählen der bis zu einem Wechselpunkte vorkommenden x eine Tafel, die folgendermaßen anfängt:

Wechselpunkte	420	460	500	540	580	620	...
$m\mathfrak{S}(x)$	0	1	2	3	5	9	...

Geht man auf diese Weise bis zu dem Wechselpunkte 780 vorwärts, so findet man an dieser Stelle das beobachtete Argument $x = 780$,

und es entsteht die Frage, ob man dieses Glied der vorangehenden oder der nachfolgenden Teilstrecke, d. h. dem Tafelargument 760 oder 800 zuweisen solle, da ja der notierte Wert 780 alle zwischen 779.5 und 780.5 liegenden Größen repräsentiert. Je nachdem das Eine oder das Andere geschieht, würde man für $m\mathfrak{S}(x)$ den Betrag 53 oder 54 erhalten. In der Regel hilft man sich nun diesem Zweifel gegenüber dadurch, daß man aus den beiden gleichrichtigen oder auch gleichunrichtigen Ansätzen das Mittel nimmt und demgemäß für $m\mathfrak{S}(x)$ den Betrag 53.5 einstellt. Der gleiche Fall würde nach Ausweis der Tabelle bei den beobachteten Argumenten 860, 900, 980, 1060 und 1100 wiederkehren.

§ 214. Das vorliegende Dilemma rührt, wie man leicht erkennt, nur davon her, daß man bei der Reduktion der vorgelegten Urliste statt des ursprünglichen Systems von Wechselpunkten ohne triftigen Grund und ganz überflüssigerweise ein völlig neues eingeführt hat, während doch der *gegebene* Weg für die Reduktion darin besteht, daß man die ursprünglichen Teilstrecken gruppenweise zusammenlegt und die hierbei entbehrlich werdenden Wechselpunkte einfach auslöscht, die übrigen aber beibehält. Überdies verstößt das erwähnte Verfahren gegen den Grundsatz, daß man beobachtete Zahlen ohne triftigen Anlaß nicht antasten soll: ein solcher Anlaß liegt aber hier nicht vor.

Nimmt man jetzt wie es sich gehört die Reduktion in der Art vor, daß man die ursprünglichen Teilstrecken gruppenweise zusammenlegt, so hat man zunächst das System der beizubehaltenden Wechselpunkte festzusetzen, weil ja bei der Zusammenlegung von je 40 Strecken ebensoviele verschiedene Wechselpunkte als Ausgangspunkt genommen werden können. Da nun die Zusammenlegung an und für sich eine willkürliche, wenn auch bei gehöriger Vorsicht unschädliche Maßregel ist, so ist die Wahl des Ausgangspunktes tatsächlich in das Belieben des Rechners gestellt; will man sich dabei an eine gewisse Regel halten, so kann man vorschreiben, daß von dem Wechselpunkte ausgegangen werde, der in der Mitte der Urliste oder doch dieser Mitte möglichst nahe liegt. Diese Regel kommt darauf hinaus, daß die äußersten Teilstrecken der reduzierten Tafel über die äußersten Glieder der Urliste an beiden Enden gleichviel hinausgreifen.

In der vorgelegten Urliste sind 428.5 und 1122.5 die von außen anliegenden Wechselpunkte, das Mittel aus beiden fällt auf den Wechselpunkt 775.5. Trägt man von dieser Stelle aus die Strecke 40 wiederholt ab, so gelangt man schließlich zu dem Wechselpunkte 415.5, der dem niedrigsten x, nämlich 429, unmittelbar vorliegt. Mit dem Punkte 415.5 lassen wir nun die reduzierte Tafel beginnen. Schreibt man für das Nummernargument X mit einer kleinen Abänderung der früher getroffenen Festsetzung jetzt stets die Zahlenreihe 0, 1, 2, . . .

vor und ordnet dem niedrigsten mitzunehmenden Wechselpunkte die Nummer $X = 0$ zu, so erhält man die Reihe der Wechselpunkte aus

$$x = 415.5 + 40X, \quad X = 0, 1, \dots 18.$$

Setzt man hiernach das neue System der Wechselpunkte an und zählt in der Urliste die bis jedem Wechselpunkt vorkommenden Glieder ab, so erhält man die Summentafel, deren Inhalt sich in gedrängter und unmittelbar verständlicher Form wie folgt schreiben läßt:

Tabelle XVIII.

$m = 217$,	$x = 415.5 + 40X$			$X = 0, \dots 18$			
$m\mathfrak{S}(x) =$	0	1	1	3	4	8	15
	21	30	50	67	89	122	153
	173	195	207	213	217	—	—

Es ist offenbar zweckmäßig, die Summentafel mit einer Null beginnen zu lassen, da man hieraus sofort ersieht, vor welchem x keine Glieder mehr vorkommen.

Für die Argumente x der Verteilungstafel gilt, da sie den Mitten der Teilstrecken entsprechen, der Ausdruck

$$x = 435.5 + 40X.$$

Bildet man also aus den Werten der Summentafel die ersten Differenzen, so erhält man die Verteilungstafel in der Gestalt

Tabelle XIX.

$m = 217$,	$x = 435.5 + 40X$,			$X = 0, \dots 17$		
$m\mathfrak{U}(x) =$	1	0	2	1	4	7
	6	9	20	17	22	33
	31	20	22	12	6	4

Das vorliegende Beispiel genügt um zu zeigen, wie die Reduktion einer Urliste angemessen durchzuführen ist.

Das benutzte Beispiel soll nun weiter dazu dienen, die Transformation des Arguments zu erläutern.

§ 215. Bei der Untersuchung eines K.-G. wird man im allgemeinen die Maßzahlen, die für das veränderliche Merkmal beobachtet worden sind, auch der Berechnung der Elemente zugrunde legen; man wird also z. B., wenn Schädelumfänge gemessen worden sind, diesen Umfang selber als Argument x wählen. Es besteht jedoch kein Hindernis statt x eine passend gewählte Funktion von x als neues Argument y einzuführen und die Elemente der nach den Werten von y konstruierten

Verteilung aufzusuchen; Gründe, die zu einer solchen Transformation Anlaß geben können, werden wir weiterhin kennen lernen.

Die zu untersuchende Transformation werde in der Gestalt $y = f(x)$ angesetzt, wo $f(x)$ vorläufig nur der einen Bedingung unterworfen sein soll, daß bei stetig wachsendem x auch y stetig wächst, daß also zu jedem x nur ein y gehört und umgekehrt. Ferner sollen für die von der Verteilung abhängenden Funktionen wie bisher die Zeichen $\mathfrak{S}$, $\mathfrak{U}$ und $\mathfrak{V}$ dienen, da z. B. in $\mathfrak{U}(x)$ und $\mathfrak{U}(y)$ das beigefügte Argument sofort erkennen läßt, ob die Verteilung nach x oder die nach y gemeint ist.

In der Urliste ist die Menge der Glieder, die mit ihrem ursprünglichen Argument unterhalb eines bestimmten x liegen, jedesmal gleich der Menge der Glieder, die mit dem transformierten Argument unterhalb des zu dem betrachteten x gehörigen y bleiben, d. h. es ist

$$\mathfrak{S}(x) = \mathfrak{S}(y). \tag{1}$$

Konstruiert man hiernach für einen unstetigen K.-G. die beiden zu x und y gehörigen Summentreppen, so besitzen die einander entsprechenden Treppenstufen gleiche Höhe, wogegen ihre Breite im allgemeinen verschieden ist. Aus (1) folgt demnach, da die Größen $\mathfrak{U}(x)$ und $\mathfrak{U}(y)$ den Höhenzuwachs von Stufe zu Stufe bedeuten, die Gleichung

$$\mathfrak{U}(x) = \mathfrak{U}(y). \tag{2}$$

Man hat also, um die Verteilung für y zu erhalten, nur nötig in der Tafel für $\mathfrak{U}(x)$ die x-Werte durch die Werte von $y = f(x)$ zu ersetzen und die $\mathfrak{U}(x)$ nunmehr als die $\mathfrak{U}(y)$ aufzufassen. Infolgedessen bietet es auch, von der Mühe der neuen Rechnung abgesehen, keinerlei Schwierigkeit, aus der neuen Tafel die numerischen Elemente von y abzuleiten.

Bei einem stetigen K.-G. folgt aus (1), da die Funktionen $\mathfrak{V}$ die Ableitung der Funktionen $\mathfrak{S}$ sind, zunächst die Gleichung

$$\mathfrak{V}(x)\,dx = \mathfrak{V}(y)\,dy, \tag{3}$$

die, wenn $g(x)$ die Ableitung von $f(x)$ bedeutet, in

$$\mathfrak{V}(x) = \mathfrak{V}(y)\,g(x) \quad \text{oder} \quad \mathfrak{V}(y) = \mathfrak{V}(x) : g(x) \tag{4}$$

übergeht. Nun werden die Elemente von y erhalten, wenn man gewisse Polynome $T(y)$, die der Reihe

$$y, \quad (y-c)^2, \quad \mathfrak{R}[h(y-c)]_2$$

angehören, der $\mathfrak{D}$-Operation unterwirft, also die zwischen den Grenzen $\pm\infty$ genommenen Integrale

$$\mathfrak{D}[T(y)] = \int T(y)\,V(y)\,dy$$

bildet, die wegen (3) in

$$\mathfrak{D}[T(y)] = \int T(y)\,\mathfrak{V}(x)\,d(x) \tag{5}$$

übergehen. Die vorstehende Gleichung läßt erkennen, daß man im allgemeinen darauf verzichten muß, den Übergang von x auf y in den Elementen durch geschlossene analytische Ausdrücke zu bewerkstelligen, denn dazu muß man erstens den analytischen Ausdruck von $\mathfrak{B}(x)$ kennen und zweitens die in (5) geforderten Integrationen auszuführen im Stande sein. Hingegen bietet der Übergang auf rein numerischem Wege an und für sich keine besonderen Schwierigkeiten. Man setzt zu dem Ende für y ein passendes System von Wechselpunkten an, berechnet die dazu gehörigen Werte von x und leitet zu diesen x die Werte von $\mathfrak{S}(x)$ vermittelst der mit den x-Elementen aufgestellten Φ-Reihe ab. Die $\mathfrak{S}(x)$ liefern dann wegen (1) die $\mathfrak{S}(y)$ für die angesetzten Wechselpunkte, so daß eine Summentafel und daraus eine Verteilungstafel entsteht, die nach dem früher entwickelten Verfahren zu den y-Elementen führt.

Das beschriebene Verfahren ist ganz zweckmäßig, wenn für $\mathfrak{S}(x)$ bereits eine befriedigende interpolatorische Darstellung durch die Φ-Reihe vorliegt, und wenn nun nachträglich noch der Übergang auf $\mathfrak{S}(y)$ bewerkstelligt werden soll. Zielt dagegen die Absicht von vornherein auf die Verteilung von y, so muß man wünschen, daß zwischen der gesuchten Verteilung und den beobachteten x ein möglichst direkter Zusammenhang hergestellt werde. Das läßt sich, wie jetzt gezeigt werden soll, ohne Schwierigkeit erreichen, sobald die Urliste mit einer genügend kleinen Abrundung aufgestellt worden ist, wie sie z. B. die Tabelle XVII aufweist. Um die Vorstellung zu fixieren, wollen wir hierbei annehmen, daß $f(x)$ den gemeinen Logarithmus $\operatorname{Log} x$ bedeute, daß es sich also nach *Fechners* Ausdruck um die Herstellung einer *logarithmischen* Verteilung handle.

§ 216. Um die verlangte Transformation auszuführen, denke man sich in Tabelle XVII zunächst jedes x durch seinen Logarithmus ersetzt und zwar mit vier Stellen, da das höchste x vier geltende Ziffern enthält. Weiter stelle man sich vor, daß die angesetzten Logarithmen durch unmittelbare Messung erhalten worden seien, wie das z. B. möglich ist, wenn bei der Messung statt eines gewöhnlichen Maßstabes mit äquidistanten Teilstrichen eine Skala verwendet wird, die nach der Art des vielverbreiteten Rechenschiebers eine logarithmische Teilung trägt. Dann läßt sich die neue Liste ebenso wie die ursprüngliche behandeln. Da die umgeformte Liste mit den Zahlen

$$2.6325 \quad 2.6964 \quad \ldots \quad 3.0492 \quad 3.0500$$

beginnt und schließt, so ist die Teilstreckenlänge gleich 0.0001, ferner lautet die Reihe der Wechselpunkte nunmehr

$$2.63245 \quad 2.63255 \quad \ldots \quad 3.04995 \quad 3.05005.$$

Die Mitte der außen anliegenden Wechselpunkte fällt auf 2.84125,

ihr Abstand ist 0.4176. Hiernach kann man unbedenklich je 300 Teilstrecken zusammenlegen und das neue System der Wechselpunkte unter Benutzung des Nummernarguments Y in der Gestalt

$$y = 2.63125 + 0.03\,Y, \qquad Y = 0, 1, \ldots 14$$

ansetzen. Schreibt man die vorstehenden y vollständig hin und schlägt dazu die Numeri auf, die wir mit x' bezeichnen wollen, so erhält man folgende Reihe:

Tabelle XX.

$x' =$	427.8	458.4	491.2	526.3	564.0
	604.3	647.5	693.8	743.4	796.6
	853.6	914.6	980.1	1050.1	1125.3

Nunmehr ist in Tabelle XVII abzuzählen, wie viele Glieder mit ihrem x jedesmal unterhalb eines x' liegen, und die abgezählte Gliedermenge als $\mathfrak{S}(x')$ oder $\mathfrak{S}(y)$ anzusetzen. Die Abzählung geht im vorliegenden Falle glatt von statten, denn die Frage, ob etwa bei $x' = 564.0$ der Gliederinhalt des Arguments $x = 564$ in die Strecke unterhalb oder oberhalb x' gehört, tritt im vorliegenden Falle nicht auf, da in Tabelle XVII das Argument 564 überhaupt nicht vorkommt.

Das angegebene Verfahren sieht auf den ersten Anblick ganz plausibel aus, ist jedoch zwei Einwendungen ausgesetzt: erstens nämlich ist die logarithmische Messung nicht wirklich erfolgt, sondern nur fingiert, zweitens wird die Tatsache außer acht gelassen, daß jedes Argument in Tabelle XVII nicht nur den hingeschriebenen Wert, sondern wegen der Abrundung alle Größen innerhalb der einschließenden Teilstrecke vertritt. So gehört z. B. das Argument $x = 854$ einer Teilstrecke t an, die von dem Wechselpunkte $u = 853.5$ bis zu dem nächsten Wechselpunkte $v = 854.5$ reicht und eine Gliedermenge J enthält, die im vorliegenden Falle gleich Eins ist. Ferner fällt in die Teilstrecke t der Punkt $x' = 853.6$, der aus dem Wechselpunkte $y = 2.93125$ entstanden ist. Da nun der wahre Wert des notierten x sowohl der Strecke von u bis x', als auch der Strecke von x' bis v angehört haben kann, so ist es angezeigt, diesen Umstand bei der Ansetzung von $\mathfrak{S}(x')$ zu berücksichtigen, indem man Streckeninhalt J angemessen auf die beiden Stücke von t verteilt. Den Weg dazu zeigt folgende geometrische Überlegung.

Man konstruiere über der x-Achse zu sämtlichen Wechselpunkten der Tabelle XVII die Summenordinaten, so daß die Reihe der Summenpunkte entsteht, durch welche die Summenkurve hindurchzuführen ist. Bei der Kleinheit der Teilstrecken erscheint es nun als unbedenklich, die Krümmung der Summenkurve zwischen je zwei benachbarten Wechselpunkten zu vernachlässigen, d. h. den Zug der Summenlinie dadurch herzustellen, daß man die gegebenen Summenpunkte

geradlinig verbindet. In diesem Falle erhält man aber $\mathfrak{S}(x')$ oder das gesuchte $\mathfrak{S}(y)$, wenn man $\mathfrak{S}(x')$ in die Strecke zwischen dem vorangehenden $\mathfrak{S}(u)$ und dem nachfolgenden $\mathfrak{S}(v)$ geradlinig hineininterpoliert.

Wendet man die vorstehende Regel auf die Tabelle XVII an, so erhält man die gesuchte logarithmische Summentafel in der Gestalt

Tabelle XXI.

$m = 217, \quad y = 2.63125 + 0.03\,Y, \quad Y = 0, 1, \cdots 14$

$m\mathfrak{S}(y) =$	0	1	1	2	4
	7	14	21	33	60
	87.1	139	177.6	205	217

Hieraus ergibt sich die logarithmische Verteilungstafel

Tabelle XXII.

$m = 217, \quad y = 2.64625 + 0.03\,Y, \quad Y = 0, 1, \cdots 13$

$m\mathfrak{U}(y) =$	1	0	1	2	3
	7	7	12	27	27.1
	51.9	38.6	27.4	12	—

Die Vernachlässigung der Krümmung erzeugt bei der Transformation selbstverständlich Fehler, die um so größer ausfallen, je länger in der zu transformierenden Tafel die Teilstrecken sind, in die man hinein zu interpolieren hat. Wenn z. B. statt der Urliste in Tabelle XVII nur die reduzierte Summentafel der Tabelle XVIII gegeben ist, so entsteht, falls man wieder für die x' der Tabelle XX geradlinig interpoliert, statt Tabelle XXI die logarithmische Summentafel

Tabelle XXIII.

$m = 217, \quad y = 2.63125 + 0.03\,Y, \quad Y = 0, 1, \ldots 14$

$m\mathfrak{S}(y) =$	0.3	1.0	1.0	2.5	3.7
	6.9	13.6	20.8	34.0	59.0
	88.0	136.8	175.5	205.4	217.0

Die Abweichungen von XXI sind in der Mehrzahl gering, gehen indessen an zwei Stellen über zwei Einheiten, also über ein Prozent des Umfanges m hinaus. Ob solche Abweichungen für groß oder für klein zu erachten sind, hängt natürlich jedesmal von der besonderen Beschaffenheit des vorgelegten K.-G. ab; immerhin ist daraus eine neue Bestätigung des Satzes zu entnehmen, daß stärkere Abrundung niemals ein Vorteil, wohl aber häufig ein wirklicher Nachteil ist.

Da bei starker Abrundung die geradlinige Interpolation der gesuchten $\mathfrak{S}(y)$ wegen der Krümmung der Summenkurve unzulässige Fehler erzeugt, und da sich auf der anderen Seite eine Berücksichtigung der Krümmung im allgemeinen nicht ohne mehr oder minder willkürliche Annahmen durchführen läßt, so ist man in solchen Fällen genötigt, auf die Interpolation der $\mathfrak{S}(y)$ zu verzichten. Es bleibt dann, wenn man nicht auf das vorhin in § 215 angegebene Verfahren zurückgreifen will, noch der folgende Weg, den wir an Tabelle XVIII deutlich machen wollen. Man ersetzt die dort auftretenden Wechselpunkte x durch ihre Logarithmen y und faßt die zugehörigen $\mathfrak{S}(x)$ nunmehr als die $\mathfrak{S}(y)$ auf. Daraus entsteht eine logarithmische Summentafel von folgender Gestalt:

Tabelle XXIV.

$y =$	2.6186	2.6585	2.6950	2.7288	2.7600
$m\mathfrak{S}(y) =$	0	1	1	3	4
$y =$	2.7892	2.8166	2.8423	2.8666	2.8896
$m\mathfrak{S}(y) =$	8	15	21	30	50
$y =$	2.9114	2.9322	2.9521	2.9710	2.9892
$m\mathfrak{S}(y) =$	67	89	122	153	173
$y =$	3.0067	3.0235	3.0396	3.0552	—
$m\mathfrak{S}(y) =$	195	207	213	217	—

Die vorgenommene Transformation, die offenbar nur die Tafelargumente, aber nicht die $\mathfrak{S}$-Größen berührt, ist völlig einwandfrei. Dieser Vorzug wird jedoch dadurch erkauft, daß die Tafelargumente nicht mehr äquidistant sind. Man kann also nicht mehr die Methode der direkten Mittelbildung anwenden, sondern ist auf die Rechnung mit den Beobachtungsgleichungen beschränkt.

§ 217. Es bedarf keiner besonderen Darlegung, wie sich die an dem Beispiele des Logarithmus erläuterte Transformation für den Fall irgend einer anderen Transformation $y = f(x)$ gestaltet. Hingegen ist es angezeigt, noch die Gründe zu besprechen, die zu einer solchen Transformation Anlaß geben können. Hierbei will ich wieder an ein konkretes Beispiel anknüpfen.

Bei der plastischen Darstellung des menschlichen Körpers ist im allgemeinen die Tendenz vorhanden, eine bestimmte Norm oder ein bestimmtes Ideal innezuhalten, was ja nicht ausschließt, daß gelegentlich, um gewisse Ideenassoziationen zu erzeugen, bewußte Abweichungen von der Norm vorgenommen werden. Tatsächlich beruhen ja die Versuche, für die Gliederung der menschlichen Gestalt einen

sogenannten *Kanon* aufzustellen, auf der Voraussetzung, daß eine solche Norm wirklich vorhanden und angebbar sei; bekannt ist der Versuch von *Zeising*, die Proportionen jener Gliederung auf den *goldenen Schnitt* als Zahlengesetz der Schönheit zurückzuführen. Selbstverständlich kann der bündige Nachweis von dem Vorhandensein einer Norm nur durch planmäßige Messungen erbracht werden, auch wenn man nach einem, *Raffael* zugeschriebenen, Ausspruch annimmt, daß die höchsten Vertreter der Kunst „den Zirkel nicht in der Hand, sondern im Auge haben". Damit wird man zu der Frage geführt, wie derartige Messungen anzuordnen und zu behandeln sind.

Um die Vorstellung zu fixieren denken wir uns, daß an einer geeigneten Reihe von Individuen die Längen x und y von Oberarm und Unterarm gleichzeitig gemessen und notiert worden seien. Man erhält dann eine Kollektivreihe $K(x, y)$ mit den beiden Argumenten x, y und kann, sobald der Umfang der Reihe sehr groß ist, daran denken, die in der XIV. Vorlesung entwickelten Beziehungen zu verwerten, um die vermutete Abhängigkeit zwischen x und y nachzuweisen. Dieser Fall tritt aber für gewöhnlich nicht ein: vielmehr wird man in der Regel froh sein müssen, wenn die Reihe vielleicht etliche hundert Glieder umfaßt. Ein solcher Umfang reicht nun allenfalls aus, um mit den früher entwickelten Kriterien das *Fehlen* der Unabhängigkeit nachzuweisen, ist aber unzureichend, wenn man über die *Form* der Abhängigkeit direkt aus den beobachteten Zahlen Aufschluß erlangen will. Man muß sich daher nach einem anderen Wege umsehen, wobei ein Kunstgriff benutzt werden kann, der in der Ausgleichungsrechnung schon seit langer Zeit Anwendung gefunden hat.

§ 218. Man denke sich bei jedem Gliede der vorgelegten Reihe außer den beiden Argumentwerten x, y auch noch den Wert des Quotienten $z = x : y$ notiert, da ja die vermutete Abhängigkeit im vorliegenden Falle darin besteht, dem Quotienten z die Tendenz nach einem konstanten Werte hin zu erteilen. Variieren nun, wie wir für den Augenblick annehmen wollen, x und y unabhängig voneinander, so ist z ein durch Mischung aus x und y entstandenes Argument, ferner sind dann die Elemente von z in ganz bestimmter Weise von den Elementen der Verteilungen $\mathfrak{V}(x)$ und $\mathfrak{V}(y)$ abhängig, und man kann sich daraufhin vornehmen, die Elemente von z aus den beobachteten Elementen von x und y zu berechnen. Andererseits liefern die beobachteten Werte des Quotienten $x : y$ eine bestimmte beobachtete Verteilung $\mathfrak{V}(z)$ mit den dazu gehörigen Elementen. Ist nun die Streuung bei der beobachteten Verteilung $\mathfrak{V}(z)$ erheblich kleiner, als die aus $\mathfrak{V}(x)$ und $\mathfrak{V}(y)$ berechnete Streuung von z, so wird man sagen dürfen, daß z in der Tat die Tendenz gegen einen konstanten Wert hin besitze.

Die angedeutete Mischungsrechnung läßt sich durchführen, liefert

jedoch so beschwerliche Formeln, daß sie für den ernstlichen Gebrauch nicht in Betracht kommt. Günstiger liegt die Sache, wenn man für z statt des Quotienten die Verbindung $x - ay$ zugrunde legt, worin a eine Konstante bedeutet. Es lassen sich dann die früher entwickelten Mischungsformeln ohne weiteres anwenden, nur daß man jetzt die Vergleichung zwischen der beobachteten und der berechneten Streuung von z für verschiedene, probeweise eingesetzte, Werte von a auszuführen und darauf dasjenige a ausfindig zu machen hat, für welches das beobachtete $\mathrm{str}(z)$ möglichst klein wird. Dieser Weg ist also immerhin gangbar, gleichwohl aber noch recht beschwerlich. Dagegen verschwinden die Schwierigkeiten sofort, wenn man die Argumente x und y vorher logarithmisch transformiert und demgemäß mit den Größen

$$X = \mathrm{Log}\, x, \quad Y = \mathrm{Log}\, y, \quad Z = X - Y$$

arbeitet. Die Tendenz gegen einen konstanten Wert von Z kommt dann dadurch zum Ausdruck, daß die beobachtete Streuung von Z kleiner wird, als der aus der Gleichung

$$\mathrm{str}(Z)^2 = \mathrm{str}(X)^2 + \mathrm{str}(Y)^2$$

berechnete Wert.

Man erkennt aus dem vorstehenden Beispiel, daß die logarithmische Transformation sich überall da empfehlen wird, wo die Abhängigkeit auf die Konstanz einer Proportion hinausläuft. Ebenso ist der Grundgedanke des Verfahrens auch dann noch anwendbar, wenn sich die vermutete Abhängigkeit in eine Gleichung von der Gestalt

$$f(x) + g(y) = \text{constans}$$

bringen läßt, wo $f(x)$ nur von x, $g(y)$ nur von y abhängt.

§ 219. Mit den vorstehenden Bemerkungen will ich die Darstellung abbrechen, die uns von der Spielaufgabe des *Chevalier de Méré* bis zu einer Frage der organischen Morphologie geführt hat. Es hat sich dabei gezeigt, daß der Aufbau einer *allgemeinen* Formenlehre der Kollektivgegenstände sehr wohl durchführbar ist, und daß die mathematischen Hindernisse, mit denen *Fechner* sich abmühte, ohne besondere Schwierigkeiten aus dem Wege geräumt werden konnten. Die ersten Keime zu einer solchen Formenlehre lassen sich bei *Fechner* sehr weit zurück verfolgen: sie finden sich bereits in einem Vortrage über *„die mathematische Behandlung organischer Gestalten und Prozesse“* (Berichte der math.-phys. Klasse der K. Sächs. Ges. d. Wiss. 1849) und treten deutlich in mehreren späteren Arbeiten zu Tage. Vielleicht darf man hoffen, daß die Kollektivmaßlehre als Hilfsmittel der Untersuchung von Massenerscheinungen vor dem Mißgeschick bewahrt bleibe, dem die Wahrscheinlichkeitsrechnung infolge unkritischer Anwendung ihrer Lehrsätze eine Zeit lang unterworfen gewesen ist.

Abkürzungen.

W.-R. = Wahrscheinlichkeitsrechnung.
rH. = relative Häufigkeit.
K.-G. = Kollektivgegenstand.
E.-G. = Exponentialgesetz.
𝔚. = mathematische Wahrscheinlichkeit.

Namenregister.

Die Zahlen geben die Paragraphen, A den Anhang an.

Anhang: Tafel der Φ-Funktionen.

Die Anordnung der nachfolgenden Tafeln bedarf keiner besonderen Erläuterung, da sie sich dem Muster der bekannten logarithmischen und trigonometrischen Tafeln eng anschließt; jedoch ist beim Gebrauche zu beachten, daß die Funktionswerte für ein positives Argument gelten, daß also für negative Argumente bei den Funktionen mit gerader Nummer das Vorzeichen umzukehren ist.

Die Tafel für $\Phi(x)$ wurde schon vor längerer Zeit auf meine Veranlassung von *B. Kämpfe* aus der siebenstelligen Tafel in *A. Meyers* *„Vorlesungen über Wahrscheinlichkeitsrechnung, deutsch von E. Czuber“* (Leipzig 1897) berechnet, und zwar nach einem Verfahren, das in meinen *„Grundlinien des wissenschaftlichen Rechnens“* in § 41 beschrieben ist. In derselben Weise ist auch die Tafel entstanden, die *G. F. Lipps* der von ihm besorgten Ausgabe von *Fechners* *„Kollektivmaßlehre“* angefügt hat. Die hier gegebenen Werte habe ich später durch Vergleichung mit *Markoff* *„Table des valeurs de l'intégrale etc.“* (St. Petersburg 1888) geprüft. Eine übersichtliche Zusammenstellung über die Transzendente $\Phi(x)$ nebst der Literatur findet man bei *H. Opitz* *„Die Kramp-Laplacesche Transzendente und ihre Umkehrung“* (Progr. Königstädt. Realgymn. Berlin, Ostern 1900).

Über die Tafeln für die Ableitungen sind bereits in § 88 des Textes die nötigen Angaben gemacht worden.

Tafel für $\Phi(x)$.

x	0	1	2	3	4	5	6	7	8	9
0.00	0.0000	0011	0023	0034	0045	0056	0068	0079	0090	0102
01	0.0113	0124	0135	0147	0158	0169	0181	0192	0203	0214
02	0.0226	0237	0248	0259	0271	0282	0293	0305	0316	0327
03	0.0338	0350	0361	0372	0384	0395	0406	0417	0429	0440
04	0.0451	0462	0474	0485	0496	0507	0519	0530	0541	0552
05	0.0564	0575	0586	0597	0609	0620	0631	0642	0654	0665
06	0.0676	0687	0699	0710	0721	0732	0744	0755	0766	0777
07	0.0789	0800	0811	0822	0833	0845	0856	0867	0878	0890
08	0.0901	0912	0923	0934	0946	0957	0968	0979	0990	1002
09	0.1013	1024	1035	1046	1058	1069	1080	1091	1102	1113
0.10	0.1125	1136	1147	1158	1169	1180	1192	1203	1214	1225
11	0.1236	1247	1259	1270	1281	1292	1303	1314	1325	1336
12	0.1348	1359	1370	1381	1392	1403	1414	1425	1436	1448
13	0.1459	1470	1481	1492	1503	1514	1525	1536	1547	1558
14	0.1569	1581	1592	1603	1614	1625	1636	1647	1658	1669
15	0.1680	1691	1702	1713	1724	1735	1746	1757	1768	1779
16	0.1790	1801	1812	1823	1834	1845	1856	1867	1878	1889
17	0.1900	1911	1922	1933	1944	1955	1966	1977	1988	1998
18	0.2009	2020	2031	2042	2053	2064	2075	2086	2097	2108
19	0.2118	2129	2140	2151	2162	2173	2184	2194	2205	2216
0.20	0.2227	2238	2249	2260	2270	2281	2292	2303	2314	2324
21	0.2335	2346	2357	2368	2378	2389	2400	2411	2421	2432
22	0.2443	2454	2464	2475	2486	2497	2507	2518	2529	2540
23	0.2550	2561	2572	2582	2593	2604	2614	2625	2636	2646
24	0.2657	2668	2678	2689	2700	2710	2721	2731	2742	2753
25	0.2763	2774	2784	2795	2806	2816	2827	2837	2848	2858
26	0.2869	2880	2890	2901	2911	2922	2932	2943	2953	2964
27	0.2974	2985	2995	3006	3016	3027	3037	3047	3058	3068
28	0.3079	3089	3100	3110	3120	3131	3141	3152	3162	3172
29	0.3183	3193	3204	3214	3224	3235	3245	3255	3266	3276
0.30	0.3286	3297	3307	3317	3327	3338	3348	3358	3369	3379
31	0.3389	3399	3410	3420	3430	3440	3450	3461	3471	3481
32	0.3491	3501	3512	3522	3532	3542	3552	3562	3573	3583
33	0.3593	3603	3613	3623	3633	3643	3653	3663	3674	3684
34	0.3694	3704	3714	3724	3734	3744	3754	3764	3774	3784
35	0.3794	3804	3814	3824	3834	3844	3854	3864	3873	3883
36	0.3893	3903	3913	3923	3933	3943	3953	3963	3972	3982
37	0.3992	4002	4012	4022	4031	4041	4051	4061	4071	4080
38	0.4090	4100	4110	4119	4129	4139	4149	4158	4168	4178
39	0.4187	4197	4207	4216	4226	4236	4245	4255	4265	4274
0.40	0.4284	4294	4303	4313	4322	4332	4341	4351	4361	4370
41	0.4380	4389	4399	4408	4418	4427	4437	4446	4456	4465
42	0.4475	4484	4494	4503	4512	4522	4531	4541	4550	4559
43	0.4569	4578	4588	4597	4606	4616	4625	4634	4644	4653
44	0.4662	4672	4681	4690	4699	4709	4718	4727	4736	4746
45	0.4755	4764	4773	4782	4792	4801	4810	4819	4828	4837
46	0.4847	4856	4865	4874	4883	4892	4901	4910	4919	4928
47	0.4937	4946	4956	4965	4974	4983	4992	5001	5010	5019
48	0.5027	5036	5045	5054	5063	5072	5081	5090	5099	5108
49	0.5117	5126	5134	5143	5152	5161	5170	5179	5187	5196
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
51	0.5292	5301	5310	5318	5327	5336	5344	5353	5362	5370
52	0.5379	5388	5396	5405	5413	5422	5430	5439	5448	5456
53	0.5465	5473	5482	5490	5499	5507	5516	5524	5533	5541
54	0.5549	5558	5566	5575	5583	5591	5600	5608	5617	5625
55	0.5633	5642	5650	5658	5667	5675	5683	5691	5700	5708
56	0.5716	5724	5733	5741	5749	5757	5765	5774	5782	5790
57	0.5798	5806	5814	5823	5831	5839	5847	5855	5863	5871
58	0.5879	5887	5895	5903	5911	5919	5927	5935	5943	5951
59	0.5959	5967	5975	5983	5991	5999	6007	6015	6023	6031
0.60	0.6039	6046	6054	6062	6070	6078	6086	6093	6101	6109
61	0.6117	6125	6132	6140	6148	6156	6163	6171	6179	6186
62	0.6194	6202	6209	6217	6225	6232	6240	6248	6255	6263
63	0.6270	6278	6286	6293	6301	6308	6316	6323	6331	6338
64	0.6346	6353	6361	6368	6376	6383	6391	6398	6405	6413
65	0.6420	6428	6435	6442	6450	6457	6464	6472	6479	6486
66	0.6494	6501	6508	6516	6523	6530	6537	6545	6552	6559
67	0.6566	6573	6581	6588	6595	6602	6609	6616	6624	6631
68	0.6638	6645	6652	6659	6666	6673	6680	6687	6694	6701
69	0.6708	6715	6722	6729	6736	6743	6750	6757	6764	6771
0.70	0.6778	6785	6792	6799	6806	6812	6819	6826	6833	6840
71	0.6847	6853	6860	6867	6874	6881	6887	6894	6901	6908
72	0.6914	6921	6928	6934	6941	6948	6954	6961	6968	6974
73	0.6981	6988	6994	7001	7007	7014	7021	7027	7034	7040
74	0.7047	7053	7060	7066	7073	7079	7086	7092	7099	7105
75	0.7112	7118	7124	7131	7137	7144	7150	7156	7163	7169
76	0.7175	7182	7188	7194	7201	7207	7213	7219	7226	7232
77	0.7238	7244	7251	7257	7263	7269	7275	7282	7288	7294
78	0.7300	7306	7312	7318	7325	7331	7337	7343	7349	7355
79	0.7361	7367	7373	7379	7385	7391	7397	7403	7409	7415
0.80	0.7421	7427	7433	7439	7445	7451	7457	7462	7468	7474
81	0.7480	7486	7492	7498	7503	7509	7515	7521	7527	7532
82	0.7538	7544	7550	7555	7561	7567	7572	7578	7584	7590
83	0.7595	7601	7607	7612	7618	7623	7629	7635	7640	7646
84	0.7651	7657	7663	7668	7674	7679	7685	7690	7696	7701
85	0.7707	7712	7718	7723	7729	7734	7739	7745	7750	7756
86	0.7761	7766	7772	7777	7782	7788	7793	7798	7804	7809
87	0.7814	7820	7825	7830	7835	7841	7846	7851	7856	7862
88	0.7867	7872	7877	7882	7888	7893	7898	7903	7908	7913
89	0.7918	7924	7929	7934	7939	7944	7949	7954	7959	7964
0.90	0.7969	7974	7979	7984	7989	7994	7999	8004	8009	8014
91	0.8019	8024	8029	8034	8038	8043	8048	8053	8058	8063
92	0.8068	8073	8077	8082	8087	8092	8097	8101	8106	8111
93	0.8116	8120	8125	8130	8135	8139	8144	8149	8153	8158
94	0.8163	8167	8172	8177	8181	8186	8191	8195	8200	8204
95	0.8209	8213	8218	8223	8227	8232	8236	8241	8245	8250
96	0.8254	8259	8263	8268	8272	8277	8281	8285	8290	8294
97	0.8299	8303	8307	8312	8316	8321	8325	8329	8334	8338
98	0.8342	8347	8351	8355	8360	8364	8368	8372	8377	8381
99	0.8385	8389	8394	8398	8402	8406	8410	8415	8419	8423
1.00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
x	0	1	2	3	4	5	6	7	8	9

Tafel für $\Phi(x)$.

x	0	1	2	3	4	5	6	7	8	9
1.00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
01	0.8468	8472	8476	8480	8484	8488	8492	8496	8500	8504
02	0.8508	8512	8516	8520	8524	8528	8532	8536	8540	8544
03	0.8548	8552	8556	8560	8563	8567	8571	8575	8579	8583
04	0.8586	8590	8594	8598	8602	8606	8609	8613	8617	8621
05	0.8624	8628	8632	8636	8639	8643	8647	8650	8654	8658
06	0.8661	8665	8669	8672	8676	8680	8683	8687	8691	8694
07	0.8698	8701	8705	8708	8712	8716	8719	8723	8726	8730
08	0.8733	8737	8740	8744	8747	8751	8754	8758	8761	8765
09	0.8768	8771	8775	8778	8782	8785	8789	8792	8795	8799
1.10	0.8802	8805	8809	8812	8815	8819	8822	8825	8829	8832
11	0.8835	8839	8842	8845	8848	8852	8855	8858	8861	8865
12	0.8868	8871	8874	8878	8881	8884	8887	8890	8893	8897
13	0.8900	8903	8906	8909	8912	8915	8918	8922	8925	8928
14	0.8931	8934	8937	8940	8943	8946	8949	8952	8955	8958
15	0.8961	8964	8967	8970	8973	8976	8979	8982	8985	8988
16	0.8991	8994	8997	9000	9003	9006	9008	9011	9014	9017
17	0.9020	9023	9026	9029	9031	9034	9037	9040	9043	9046
18	0.9048	9051	9054	9057	9060	9062	9065	9068	9071	9073
19	0.9076	9079	9082	9084	9087	9090	9092	9095	9098	9100
1.20	0.9103	9106	9108	9111	9114	9116	9119	9122	9124	9127
21	0.9130	9132	9135	9137	9140	9143	9145	9148	9150	9153
22	0.9155	9158	9160	9163	9165	9168	9171	9173	9176	9178
23	0.9181	9183	9185	9188	9190	9193	9195	9198	9200	9203
24	0.9205	9207	9210	9212	9215	9217	9219	9222	9224	9227
25	0.9229	9231	9234	9236	9238	9241	9243	9245	9248	9250
26	0.9252	9255	9257	9259	9262	9264	9266	9268	9271	9273
27	0.9275	9277	9280	9282	9284	9286	9289	9291	9293	9295
28	0.9297	9300	9302	9304	9306	9308	9310	9313	9315	9317
29	0.9319	9321	9323	9325	9327	9330	9332	9334	9336	9338
1.30	0.9340	9342	9344	9346	9348	9350	9352	9355	9357	9359
31	0.9361	9363	9365	9367	9369	9371	9373	9375	9377	9379
32	0.9381	9383	9385	9387	9389	9390	9392	9394	9396	9398
33	0.9400	9402	9404	9406	9408	9410	9412	9413	9415	9417
34	0.9419	9421	9423	9425	9427	9428	9430	9432	9434	9436
35	0.9438	9439	9441	9443	9445	9447	9448	9450	9452	9454
36	0.9456	9457	9459	9461	9463	9464	9466	9468	9470	9471
37	0.9473	9475	9477	9478	9480	9482	9483	9485	9487	9488
38	0.9490	9492	9494	9495	9497	9499	9500	9502	9503	9505
39	0.9507	9508	9510	9512	9513	9515	9516	9518	9520	9521
1.40	0.9523	9524	9526	9528	9529	9531	9532	9534	9535	9537
41	0.9539	9540	9542	9543	9545	9546	9548	9549	9551	9552
42	0.9554	9555	9557	9558	9560	9561	9563	9564	9566	9567
43	0.9569	9570	9571	9573	9574	9576	9577	9579	9580	9582
44	0.9583	9584	9586	9587	9589	9590	9591	9593	9594	9596
45	0.9597	9598	9600	9601	9602	9604	9605	9607	9608	9609
46	0.9611	9612	9613	9615	9616	9617	9618	9620	9621	9622
47	0.9624	9625	9626	9628	9629	9630	9631	9633	9634	9635
48	0.9637	9638	9639	9640	9642	9643	9644	9645	9647	9648
49	0.9649	9650	9651	9653	9654	9655	9656	9657	9659	9660
1.50	0.9661	9662	9663	9665	9666	9667	9668	9669	9670	9672
x	0	1	2	3	4	5	6	7	8	9

Tafel für $\Phi(x)$.

x	0	1	2	3	4	5	6	7	8	9
1.5	0.9661	9673	9684	9695	9706	9716	9726	9736	9745	9755
1.6	0.9763	9772	9780	9788	9796	9804	9811	9818	9825	9832
1.7	0.9838	9844	9850	9856	9861	9867	9872	9877	9882	9886
1.8	0.9891	9895	9899	9903	9907	9911	9915	9918	9922	9925
1.9	0.9928	9931	9934	9937	9939	9942	9944	9947	9949	9951
2.0	0.9953	9955	9957	9959	9961	9963	9964	9966	9967	9969
2.1	0.9970	9972	9973	9974	9975	9976	9977	9979	9980	9980
2.2	0.9981	9982	9983	9984	9985	9985	9986	9987	9987	9988
2.3	0.9989	9989	9990	9990	9991	9991	9992	9992	9992	9993
2.4	0.9993	9993	9994	9994	9994	9995	9995	9995	9995	9996
2.5	0.9996	9996	9996	9997	9997	9997	9997	9997	9997	9998
2.6	0.9998	9998	9998	9998	9998	9998	9998	9998	9998	9999
2.7	0.9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
2.8	0.9999	9999	9999	9999	9999	9999	9999	0000	0000	0000

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
0.00	+ 1.1284	1	0.0000	113	− 0.5642	2
.01	+ 1.1283	4	− 0.0113	113	− 0.5640	5
.02	1.1279	5	0.0226	112	0.5635	8
.03	1.1274	8	0.0338	113	0.5627	12
.04	+ 1.1266	10	− 0.0451	112	− 0.5615	15
.05	1.1256	13	0.0563	112	0.5600	19
.06	1.1243	14	0.0675	111	0.5581	22
.07	+ 1.1229	17	− 0.0786	111	− 0.5559	25
.08	1.1212	19	0.0897	110	0.5534	28
.09	1.1193	21	0.1007	110	0.5506	32
0.10	+ 1.1172	24	− 0.1117	109	− 0.5474	35
.11	+ 1.1148	26	− 0.1226	109	− 0.5439	38
.12	1.1122	27	0.1335	107	0.5401	41
.13	1.1095	30	0.1442	107	0.5360	44
.14	+ 1.1065	32	− 0.1549	106	− 0.5316	48
.15	1.1033	34	0.1655	105	0.5268	50
.16	1.0999	37	0.1760	104	0.5218	54
.17	+ 1.0962	38	− 0.1864	102	− 0.5164	56
.18	1.0924	40	0.1966	102	0.5108	59
.19	1.0884	43	0.2068	100	0.5049	62
0.20	+ 1.0841	44	− 0.2168	99	− 0.4987	65
.21	+ 1.0797	46	− 0.2267	98	− 0.4922	67
.22	1.0751	49	0.2365	97	0.4855	70
.23	1.0702	50	0.2462	95	0.4785	72
.24	+ 1.0652	52	− 0.2557	93	− 0.4713	75
.25	1.0600	54	0.2650	92	0.4638	78
.26	1.0546	56	0.2742	90	0.4560	80
.27	+ 1.0490	57	− 0.2832	89	− 0.4480	81
.28	1.0433	59	0.2921	87	0.4399	85
.29	1.0374	61	0.3008	86	0.4314	86
0.30	+ 1.0313	63	− 0.3094	83	− 0.4228	88
.31	+ 1.0250	64	− 0.3177	82	− 0.4140	90
.32	1.0186	66	0.3259	80	0.4050	92
.33	1.0120	68	0.3339	79	0.3958	94
.34	+ 1.0052	69	− 0.3418	76	− 0.3864	95
.35	0.9983	71	0.3494	74	0.3769	98
.36	0.9912	72	0.3568	73	0.3671	98
.37	+ 0.9840	73	− 0.3641	70	− 0.3573	100
.38	0.9767	75	0.3711	69	0.3473	101
.39	0.9692	77	0.3780	66	0.3372	103
0.40	+ 0.9615	77	− 0.3846	65	− 0.3269	103
.41	+ 0.9538	79	− 0.3911	62	− 0.3166	105
.42	0.9459	80	0.3973	60	0.3061	106
.43	0.9379	81	0.4033	58	0.2955	106
.44	+ 0.9298	83	− 0.4091	56	− 0.2849	107
.45	0.9215	83	0.4147	54	0.2742	108
.46	0.9132	85	0.4201	51	0.2634	109
.47	+ 0.9047	85	− 0.4252	50	− 0.2525	109
.48	0.8962	87	0.4302	47	0.2416	109
.49	0.8875	87	0.4349	45	0.2307	110
0.50	+ 0.8788		− 0.4394		− 0.2197	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
0.00	0.0000		+ 0.8463		0.0000	
.01	+ 0.0169	169	+ 0.8459	4	− 0.0423	423
.02	0.0338	169	0.8446	13	0.0845	422
.03	0.0507	169	0.8425	21	0.1267	422
.04	+ 0.0675	168	+ 0.8395	30	− 0.1686	419
.05	0.0843	168	0.8357	38	0.2103	417
.06	0.1009	166	0.8311	46	0.2518	415
.07	+ 0.1175	166	+ 0.8257	54	− 0.2928	410
.08	0.1340	165	0.8194	63	0.3335	407
.09	0.1503	163	0.8123	71	0.3737	402
0.10	+ 0.1665	162	+ 0.8045	78	− 0.4134	397
.11	+ 0.1825	160	+ 0.7958	87	− 0.4525	391
.12	0.1983	158	0.7864	94	0.4909	384
.13	0.2139	156	0.7762	102	0.5287	378
.14	+ 0.2293	154	+ 0.7652	110	− 0.5658	371
.15	0.2445	152	0.7535	117	0.6021	363
.16	0.2595	150	0.7411	124	0.6375	354
.17	+ 0.2742	147	+ 0.7280	131	− 0.6721	346
.18	0.2886	144	0.7143	137	0.7057	336
.19	0.3027	141	0.6998	145	0.7384	327
0.20	+ 0.3166	139	+ 0.6847	151	− 0.7701	317
.21	+ 0.3301	135	+ 0.6690	157	− 0.8007	306
.22	0.3433	132	0.6527	163	0.8302	295
.23	0.3562	129	0.6358	169	0.8587	285
.24	+ 0.3688	126	+ 0.6184	174	− 0.8859	272
.25	0.3809	121	0.6004	180	0.9120	261
.26	0.3928	119	0.5819	185	0.9368	248
.27	+ 0.4042	114	+ 0.5629	190	− 0.9604	236
.28	0.4153	111	0.5435	194	0.9827	223
.29	0.4260	107	0.5236	199	1.0038	211
0.30	+ 0.4362	102	+ 0.5034	202	− 1.0235	197
.31	+ 0.4461	99	+ 0.4827	207	− 1.0418	183
.32	0.4555	94	0.4617	210	1.0588	170
.33	0.4646	91	0.4404	213	1.0744	156
.34	+ 0.4731	85	+ 0.4187	217	− 1.0887	143
.35	0.4813	82	0.3968	219	1.1015	128
.36	0.4890	77	0.3747	221	1.1129	114
.37	+ 0.4963	73	+ 0.3523	224	− 1.1229	100
.38	0.5031	68	0.3298	225	1.1315	86
.39	0.5095	64	0.3071	227	1.1387	72
0.40	+ 0.5154	59	+ 0.2842	229	− 1.1445	58
.41	+ 0.5208	54	+ 0.2613	229	− 1.1488	43
.42	0.5258	50	0.2383	230	1.1518	30
.43	0.5304	46	0.2152	231	1.1533	15
.44	+ 0.5344	40	+ 0.1922	230	− 1.1534	1
.45	0.5381	37	0.1691	231	1.1522	12
.46	0.5412	31	0.1461	230	1.1496	26
.47	+ 0.5439	27	+ 0.1231	230	− 1.1457	39
.48	0.5461	22	0.1003	228	1.1404	53
.49	0.5479	18	0.0775	228	1.1338	66
0.50	+ 0.5492	13	+ 0.0549	226	− 1.1259	79

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
0.50	+ 0.8788	88	− 0.4394	43	− 0.2197	110
.51	+ 0.8700	90	− 0.4437	40	− 0.2087	110
.52	0.8610	90	0.4477	39	0.1977	110
.53	0.8520	90	0.4516	36	0.1867	110
.54	+ 0.8430	92	− 0.4552	34	− 0.1757	110
.55	0.8338	92	0.4586	32	0.1647	110
.56	0.8246	92	0.4618	30	0.1537	109
.57	+ 0.8154	94	− 0.4648	27	− 0.1428	109
.58	0.8060	93	0.4675	25	0.1319	109
.59	0.7967	95	0.4700	23	0.1210	108
0.60	+ 0.7872	94	− 0.4723	21	− 0.1102	107
.61	+ 0.7778	95	− 0.4744	19	− 0.0995	107
.62	0.7683	96	0.4763	17	0.0888	106
.63	0.7587	96	0.4780	15	0.0782	105
.64	+ 0.7491	96	− 0.4795	12	− 0.0677	104
.65	0.7395	96	0.4807	10	0.0573	103
.66	0.7299	96	0.4817	9	0.0470	102
.67	+ 0.7203	97	− 0.4826	6	− 0.0368	101
.68	0.7106	96	0.4832	5	0.0267	99
.69	0.7010	97	0.4837	2	0.0168	99
0.70	+ 0.6913	97	− 0.4839	0	− 0.0069	97
.71	+ 0.6816	97	− 0.4839	1	+ 0.0028	96
.72	0.6719	97	0.4838	4	0.0124	94
.73	0.6622	96	0.4834	5	0.0218	93
.74	+ 0.6526	97	− 0.4829	7	+ 0.0311	91
.75	0.6429	96	0.4822	9	0.0402	89
.76	0.6333	96	0.4813	11	0.0491	88
.77	+ 0.6237	96	− 0.4802	12	+ 0.0579	87
.78	0.6141	96	0.4790	14	0.0666	84
.79	0.6045	95	0.4776	16	0.0750	83
0.80	+ 0.5950	95	− 0.4760	18	+ 0.0833	81
.81	+ 0.5855	95	− 0.4742	19	+ 0.0914	79
.82	0.5760	94	0.4723	20	0.0993	77
.83	0.5666	94	0.4703	22	0.1070	76
.84	+ 0.5572	93	− 0.4681	24	+ 0.1146	73
.85	0.5479	93	0.4657	25	0.1219	71
.86	0.5386	93	0.4632	27	0.1290	70
.87	+ 0.5293	91	− 0.4605	28	+ 0.1360	67
.88	0.5202	92	0.4577	29	0.1427	66
.89	0.5110	90	0.4548	30	0.1493	63
0.90	+ 0.5020	90	− 0.4518	32	+ 0.1556	61
.91	+ 0.4930	90	− 0.4486	33	+ 0.1617	60
.92	0.4840	88	0.4453	34	0.1677	57
.93	0.4752	88	0.4419	35	0.1734	55
.94	+ 0.4664	88	− 0.4384	37	+ 0.1789	53
.95	0.4576	86	0.4347	37	0.1842	51
.96	0.4490	86	0.4310	38	0.1893	49
.97	+ 0.4404	85	− 0.4272	40	+ 0.1942	46
.98	0.4319	84	0.4232	40	0.1988	45
.99	0.4235	84	0.4192	41	0.2033	43
1.00	+ 0.4151		− 0.4151		+ 0.2076	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
0.50	+ 0.5492	9	+ 0.0549	224	— 1.1259	91
.51	+ 0.5501	4	+ 0.0325	222	— 1.1168	104
.52	0.5505	0	+ 0.0103	221	1.1064	116
.53	0.5505	4	— 0.0118	217	1.0948	128
.54	+ 0.5501	9	— 0.0335	215	— 1.0820	139
.55	0.5492	13	0.0550	212	1.0681	151
.56	0.5479	18	0.0762	209	1.0530	161
.57	+ 0.5461	21	— 0.0971	206	— 1.0369	172
.58	0.5440	26	0.1177	202	1.0197	182
.59	0.5414	29	0.1379	199	1.0015	192
0.60	+ 0.5385	34	— 0.1578	194	— 0.9823	202
.61	+ 0.5351	37	— 0.1772	190	— 0.9621	210
.62	0.5314	41	0.1962	186	0.9411	219
.63	0.5273	45	0.2148	182	0.9192	227
.64	+ 0.5228	48	— 0.2330	177	— 0.8965	235
.65	0.5180	52	0.2507	172	0.8730	243
.66	0.5128	56	0.2679	167	0.8487	249
.67	+ 0.5072	58	— 0.2846	163	— 0.8238	256
.68	0.5014	62	0.3009	157	0.7982	262
.69	0.4952	65	0.3166	151	0.7720	268
0.70	+ 0.4887	68	— 0.3317	147	— 0.7452	272
.71	+ 0.4819	70	— 0.3464	141	— 0.7180	278
.72	0.4749	74	0.3605	135	0.6902	281
.73	0.4675	76	0.3740	129	0.6621	286
.74	+ 0.4599	78	— 0.3869	124	— 0.6335	289
.75	0.4521	81	0.3993	118	0.6046	291
.76	0.4440	84	0.4111	112	0.5755	295
.77	+ 0.4356	85	— 0.4223	107	— 0.5460	296
.78	0.4271	88	0.4330	100	0.5164	298
.79	0.4183	89	0.4430	94	0.4866	298
0.80	+ 0.4094	92	— 0.4524	89	— 0.4568	300
.81	+ 0.4002	93	— 0.4613	82	— 0.4268	300
.82	0.3909	95	0.4695	76	0.3968	299
.83	0.3814	96	0.4771	71	0.3669	300
.84	+ 0.3718	97	— 0.4842	64	— 0.3369	298
.85	0.3621	99	0.4906	59	0.3071	297
.86	0.3522	100	0.4965	52	0.2774	295
.87	+ 0.3422	101	— 0.5017	47	— 0.2479	292
.88	0.3321	101	0.5064	41	0.2187	291
.89	0.3220	103	0.5105	35	0.1896	287
0.90	+ 0.3117	103	— 0.5140	29	— 0.1609	285
.91	+ 0.3014	103	— 0.5169	24	— 0.1324	280
.92	0.2911	105	0.5193	18	0.1044	277
.93	0.2806	104	0.5211	12	0.0767	273
.94	+ 0.2702	104	— 0.5223	8	— 0.0494	268
.95	0.2598	105	0.5231	1	— 0.0226	263
.96	0.2493	105	0.5232	3	+ 0.0037	259
.97	+ 0.2388	104	— 0.5229	8	+ 0.0296	253
.98	0.2284	104	0.5221	14	0.0549	247
.99	0.2180	104	0.5207	18	0.0796	242
1.00	+ 0.2076		— 0.5189		+ 0.1038	

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
1.00	+ 0.4151	83	− 0.4151	42	+ 0.2076	40
.01	+ 0.4068	81	− 0.4109	43	+ 0.2116	38
.02	0.3987	81	0.4066	43	0.2154	37
.03	0.3906	80	0.4023	44	0.2191	34
.04	+ 0.3826	79	− 0.3979	45	+ 0.2225	32
.05	0.3747	79	0.3934	45	0.2257	31
.06	0.3668	77	0.3889	46	0.2288	28
.07	+ 0.3591	76	− 0.3843	47	+ 0.2316	26
.08	0.3515	76	0.3796	47	0.2342	25
.09	0.3439	74	0.3749	48	0.2367	22
1.10	+ 0.3365	74	− 0.3701	48	+ 0.2389	21
.11	+ 0.3291	72	− 0.3653	48	+ 0.2410	18
.12	0.3219	72	0.3605	49	0.2428	17
.13	0.3147	71	0.3556	49	0.2445	15
.14	+ 0.3076	69	− 0.3507	49	+ 0.2460	13
.15	0.3007	69	0.3458	50	0.2473	11
.16	0.2938	68	0.3408	50	0.2484	10
.17	+ 0.2870	66	− 0.3358	50	+ 0.2494	8
.18	0.2804	66	0.3308	50	0.2502	6
.19	0.2738	65	0.3258	50	0.2508	5
1.20	+ 0.2673	63	− 0.3208	50	+ 0.2513	3
.21	+ 0.2610	63	− 0.3158	51	+ 0.2516	2
.22	0.2547	62	0.3107	50	0.2518	0
.23	0.2485	60	0.3057	50	0.2518	2
.24	+ 0.2425	60	− 0.3007	50	+ 0.2516	3
.25	0.2365	58	0.2957	51	0.2513	4
.26	0.2307	58	0.2906	50	0.2509	6
.27	+ 0.2249	57	− 0.2856	50	+ 0.2503	7
.28	0.2192	55	0.2806	50	0.2496	9
.29	0.2137	55	0.2756	49	0.2487	9
1.30	+ 0.2082	54	− 0.2707	50	+ 0.2478	11
.31	+ 0.2028	52	− 0.2657	49	+ 0.2467	12
.32	0.1976	52	0.2608	49	0.2455	13
.33	0.1924	51	0.2559	49	0.2442	15
.34	+ 0.1873	49	− 0.2510	48	+ 0.2427	15
.35	0.1824	49	0.2462	48	0.2412	17
.36	0.1775	48	0.2414	48	0.2395	17
.37	+ 0.1727	47	− 0.2366	47	+ 0.2378	18
.38	0.1680	46	0.2319	47	0.2360	19
.39	0.1634	45	0.2272	47	0.2341	20
1.40	+ 0.1589	44	− 0.2225	46	+ 0.2321	21
.41	+ 0.1545	43	− 0.2179	46	+ 0.2300	22
.42	0.1502	42	0.2133	45	0.2278	22
.43	0.1460	41	0.2088	45	0.2256	23
.44	+ 0.1419	41	− 0.2043	44	+ 0.2233	24
.45	0.1378	39	0.1999	44	0.2209	25
.46	0.1339	39	0.1955	44	0.2184	25
.47	+ 0.1300	38	− 0.1911	43	+ 0.2159	25
.48	0.1262	37	0.1868	42	0.2134	26
.49	0.1225	36	0.1826	42	0.2108	27
1.50	+ 0.1189		− 0.1784		+ 0.2081	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
1.00	+ 0.2076	104	— 0.5189	23	+ 0.1038	235
.01	+ 0.1972	103	— 0.5166	28	+ 0.1273	230
.02	0.1869	103	0.5138	32	0.1503	223
.03	0.1766	101	0.5106	37	0.1726	216
.04	+ 0.1665	101	— 0.5069	41	+ 0.1942	210
.05	0.1564	100	0.5028	45	0.2152	203
.06	0.1464	100	0.4983	49	0.2355	195
.07	+ 0.1364	98	— 0.4934	53	+ 0.2550	189
.08	0.1266	97	0.4881	57	0.2739	181
.09	0.1169	96	0.4824	60	0.2920	174
1.10	+ 0.1073	94	— 0.4764	63	+ 0.3094	166
.11	+ 0.0979	94	— 0.4701	67	+ 0.3260	159
.12	0.0885	92	0.4634	70	0.3419	151
.13	0.0793	90	0.4564	73	0.3570	144
.14	+ 0.0703	89	— 0.4491	76	+ 0.3714	136
.15	0.0614	88	0.4415	78	0.3850	129
.16	0.0526	86	0.4337	81	0.3979	120
.17	+ 0.0440	84	— 0.4256	83	+ 0.4099	113
.18	0.0356	83	0.4173	85	0.4212	106
.19	0.0273	81	0.4088	87	0.4318	98
1.20	+ 0.0192	79	— 0.4001	90	+ 0.4416	90
.21	+ 0.0113	77	— 0.3911	91	+ 0.4506	83
.22	+ 0.0036	75	0.3820	92	0.4589	75
.23	— 0.0039	74	0.3728	94	0.4664	68
.24	— 0.0113	72	— 0.3634	95	+ 0.4732	61
.25	0.0185	70	0.3539	97	0.4793	53
.26	0.0255	67	0.3442	97	0.4846	47
.27	— 0.0322	66	— 0.3345	98	+ 0.4893	39
.28	0.0388	64	0.3247	99	0.4932	33
.29	0.0452	62	0.3148	100	0.4965	26
1.30	— 0.0514	60	— 0.3048	100	+ 0.4991	19
.31	— 0.0574	58	— 0.2948	100	+ 0.5010	13
.32	0.0632	56	0.2848	101	0.5023	7
.33	0.0688	54	0.2747	101	0.5030	0
.34	— 0.0742	52	— 0.2646	100	+ 0.5030	5
.35	0.0794	50	0.2546	101	0.5025	11
.36	0.0844	48	0.2445	100	0.5014	17
.37	— 0.0892	46	— 0.2345	99	+ 0.4997	23
.38	0.0938	44	0.2246	100	0.4974	27
.39	0.0982	42	0.2146	98	0.4947	33
1.40	— 0.1024	40	— 0.2048	98	+ 0.4914	38
.41	— 0.1064	38	— 0.1950	97	+ 0.4876	42
.42	0.1102	36	0.1853	96	0.4834	47
.43	0.1138	34	0.1757	96	0.4787	51
.44	— 0.1172	32	— 0.1661	94	+ 0.4736	55
.45	0.1204	31	0.1567	93	0.4681	60
.46	0.1235	28	0.1474	92	0.4621	63
.47	— 0.1263	27	— 0.1382	90	+ 0.4558	66
.48	0.1290	25	0.1292	89	0.4492	70
.49	0.1315	23	0.1203	88	0.4422	74
1.50	— 0.1338		— 0.1115		+ 0.4348	

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$		$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
1.50	+ 0.1189	35	— 0.1784	41	+ 0.2081	27
.51	+ 0.1154	34	— 0.1743	41	+ 0.2054	27
.52	0.1120	34	0.1702	40	0.2027	28
.53	0.1086	33	0.1662	40	0.1999	28
.54	+ 0.1053	32	— 0.1622	39	+ 0.1971	28
.55	0.1021	31	0.1583	39	0 1943	29
.56	0.0990	31	0.1544	38	0.1914	29
.57	+ 0.0959	29	— 0.1506	37	+ 0.1885	29
.58	0.0930	29	0.1469	37	0.1856	30
.59	0.0901	29	0.1432	36	0.1826	29
1.60	+ 0.0872	27	— 0.1396	36	+ 0.1797	30
.61	+ 0.0845	27	— 0.1360	35	+ 0.1767	29
.62	0.0818	26	0.1325	34	0.1738	30
.63	0.0792	26	0.1291	34	0.1708	30
.64	+ 0.0766	25	— 0.1257	34	+ 0.1678	30
.65	0.0741	24	0.1223	32	0.1648	30
.66	0.0717	23	0.1191	32	0.1618	30
.67	+ 0.0694	23	— 0.1159	32	+ 0.1588	30
.68	0.0671	22	0.1127	31	0.1558	30
.69	0.0649	22	0.1096	30	0.1528	29
1.70	+ 0.0627	21	— 0.1066	30	+ 0.1499	30
.71	+ 0.0606	20	— 0.1036	29	+ 0.1469	29
.72	0.0586	20	0.1007	28	0.1440	30
.73	0.0566	20	0.0979	28	0.1410	29
.74	+ 0.0546	18	— 0.0951	27	+ 0.1381	29
.75	0.0528	18	0.0924	27	0.1352	28
.76	0.0510	18	0.0897	26	0.1324	29
.77	+ 0.0492	17	— 0.0871	26	+ 0.1295	28
.78	0.0475	17	0.0845	25	0.1267	28
.79	0.0458	16	0.0820	25	0.1239	28
1.80	+ 0.0442	16	— 0.0795	23	+ 0.1211	28
.81	+ 0.0426	15	— 0.0772	24	+ 0.1183	27
.82	0.0411	15	0.0748	23	0.1156	27
.83	0.0396	14	0.0725	22	0.1129	27
.84	+ 0.0382	14	— 0.0703	22	+ 0.1102	26
.85	0.0368	13	0.0681	21	0.1076	26
.86	0.0355	13	0.0660	21	0.1050	26
.87	+ 0.0342	13	— 0.0639	20	+ 0.1024	25
.88	0.0329	12	0.0619	20	0.0999	25
.89	0.0317	12	0.0599	19	0.0974	25
1.90	+ 0.0305	11	— 0.0580	19	+ 0.0949	24
.91	+ 0.0294	11	— 0.0561	18	+ 0.0925	24
.92	0.0283	11	0.0543	18	0.0901	23
.93	0.0272	10	0.0525	17	0.0878	24
.94	+ 0.0262	10	— 0.0508	17	+ 0.0854	22
.95	0.0252	10	0.0491	16	0.0832	23
.96	0.0242	9	0.0475	16	0.0809	22
.97	+ 0.0233	9	— 0.0459	16	+ 0.0787	22
.98	0.0224	9	0.0443	15	0.0765	21
.99	0.0215	8	0.0428	15	0.0744	21
2.00	+ 0.0207		— 0.0413		+ 0.0723	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
1.50	− 0.1338	21	− 0.1115	86	+ 0.4348	76
.51	− 0.1359	20	− 0.1029	85	+ 0.4272	79
.52	0.1379	18	0.0944	83	0.4193	81
.53	0.1397	17	0.0861	81	0.4112	84
.54	− 0 1414	14	− 0.0780	80	+ 0.4028	86
.55	0.1428	14	0.0700	78	0.3942	89
.56	0.1442	11	0.0622	76	0.3853	90
.57	− 0.1453	10	− 0.0546	75	+ 0.3763	91
.58	0.1463	9	0.0471	72	0.3672	93
.59	0.1472	7	0.0399	71	0.3579	95
1.60	− 0.1479	6	− 0.0328	68	+ 0.3484	95
.61	− 0.1485	5	− 0.0260	67	+ 0.3389	97
.62	0.1490	3	0.0193	65	0.3292	97
.63	0.1493	2	0.0128	63	0.3195	99
.64	− 0.1495	1	− 0.0065	61	+ 0.3096	98
.65	0.1496	1	− 0.0004	59	0.2998	99
.66	0.1495	2	+ 0.0055	57	0.2899	99
.67	− 0.1493	2	+ 0.0112	55	+ 0.2800	99
.68	0.1491	4	0.0167	53	0.2701	99
.69	0.1487	5	0.0220	51	0.2602	99
1.70	− 0.1482	6	+ 0.0271	49	+ 0.2503	98
.71	− 0.1476	7	+ 0.0320	47	+ 0.2405	98
.72	0.1469	8	0.0367	45	0.2307	98
.73	0.1461	8	0.0412	44	0.2209	96
.74	− 0.1453	10	+ 0.0456	41	+ 0.2113	96
.75	0.1443	10	0.0497	39	0.2017	95
.76	0.1433	11	0.0536	38	0.1922	94
.77	− 0.1422	12	+ 0.0574	35	+ 0.1828	93
.78	0.1410	13	0.0609	34	0.1735	92
.79	0.1397	13	0.0643	32	0.1643	90
1.80	− 0.1384	14	+ 0.0675	30	+ 0.1553	89
.81	− 0.1370	14	+ 0.0705	29	+ 0.1464	87
.82	0.1356	15	0.0734	26	0.1377	86
.83	0.1341	16	0.0760	25	0.1291	85
.84	− 0.1325	15	+ 0.0785	24	+ 0.1206	83
.85	0.1310	17	0.0809	21	0.1123	81
.86	0.1293	17	0.0830	20	0.1042	79
.87	− 0.1276	17	+ 0.0850	19	+ 0.0963	78
.88	0.1259	17	0.0869	17	0.0885	76
.89	0.1242	18	0.0886	15	0.0809	74
1.90	− 0.1224	18	+ 0.0901	14	+ 0.0735	72
.91	− 0.1206	19	+ 0.0915	13	+ 0.0663	70
.92	0.1187	19	0.0928	11	0.0593	68
.93	0.1168	18	0.0939	10	0.0525	66
.94	− 0.1150	19	+ 0.0949	8	+ 0.0459	64
.95	0.1131	20	0.0957	7	0.0395	63
.96	0.1111	19	0.0964	7	0.0332	60
.97	− 0.1092	19	+ 0.0971	4	+ 0.0272	58
.98	0.1073	20	0.0975	4	0.0214	56
.99	0.1053	20	0.0979	3	0.0158	55
2.00	− 0 1033		+ 0.0982		+ 0.0103	

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$	$\Phi(x)_2 : 2$		$\Phi(x)_3 : 4$	
2.00	+ 0.0207	− 0.0413	14	+ 0.0723	20
.01	+ 0.0199	− 0.0399	14	+ 0.0703	20
.02	0.0191	0.0385	13	0.0683	20
.03	0.0183	0.0372	13	0.0663	19
.04	+ 0.0176	− 0.0359	13	+ 0.0644	19
.05	0.0169	0.0346	12	0 0625	19
.06	0.0162	0.0334	12	0.0606	18
.07	+ 0.0155	− 0.0322	12	+ 0.0588	17
.08	0.0149	0.0310	11	0.0571	18
.09	0.0143	0.0299	11	0.0553	17
2.10	+ 0.0137	− 0.0288	11	+ 0.0536	16
.11	+ 0.0132	− 0.0277	10	+ 0.0520	16
.12	0.0126	0.0267	10	0.0504	16
.13	0.0121	0.0257	9	0.0488	16
.14	+ 0.0116	− 0.0248	10	+ 0.0472	15
.15	0.0111	0.0238	9	0 0457	15
.16	0.0106	0.0229	8	0.0442	14
.17	+ 0.0102	− 0.0221	9	+ 0.0428	14
.18	0.0097	0.0212	8	0.0414	13
.19	0.0093	0.0204	8	0.0401	14
2.20	+ 0.0089	− 0.0196	7	+ 0.0387	13
.21	+ 0.0085	− 0.0189	8	+ 0.0374	12
.22	0.0082	0.0181	7	0.0362	13
.23	0.0078	0.0174	7	0.0349	12
.24	+ 0.0075	− 0.0167	6	+ 0.0337	11
.25	0.0071	0.0161	7	0.0326	11
.26	0.0068	0.0154	6	0.0315	11
.27	+ 0.0065	− 0.0148	6	+ 0.0304	11
.28	0.0062	0.0142	6	0.0293	10
.29	0.0060	0.0136	5	0.0283	10
2.30	+ 0.0057	− 0.0131	6	+ 0.0273	10
.31	+ 0.0054	− 0.0125	5	+ 0.0263	10
.32	0.0052	0.0120	5	0.0253	9
.33	0.0050	0.0115	4	0.0244	9
34	+ 0.0047	− 0.0111	5	+ 0.0235	9
.35	0.0045	0.0106	4	0.0226	8
.36	0.0043	0.0102	5	0.0218	8
.37	+ 0.0041	− 0.0097	4	+ 0.0210	8
.38	0.0039	0.0093	4	0.0202	8
.39	0.0037	0.0089	4	0.0194	7
2.40	+ 0.0036	− 0.0085	3	+ 0.0187	7
.41	+ 0.0034	− 0.0082	4	+ 0.0180	7
.42	0.0032	0.0078	3	0.0173	7
.43	0.0031	0.0075	4	0.0166	6
.44	+ 0.0029	− 0.0071	3	+ 0.0160	6
.45	0.0028	0.0068	3	0.0154	7
.46	0.0027	0.0065	3	0.0147	5
.47	+ 0.0025	− 0.0062	2	+ 0.0142	6
.48	0.0024	0.0060	3	0.0136	5
.49	0.0023	0.0057	3	0.0131	6
2.50	+ 0.0022	− 0.0054		+ 0.0125	

x	$\Phi(x)_4 : 8$		$\Phi(x)_5 : 16$		$\Phi(x)_6 : 32$	
2.00	− 0.1033	19	+ 0.0982	1	+ 0.0103	52
.01	− 0.1014	20	+ 0.0983	1	+ 0.0051	50
.02	0.0994	20	0.0984	1	+ 0.0001	48
.03	0.0974	19	0.0983	1	− 0.0047	47
.04	− 0.0955	20	+ 0.0982	2	− 0.0094	44
.05	0.0935	19	0.0980	4	0.0138	42
.06	0.0916	20	0.0976	4	0.0180	41
.07	− 0.0896	19	+ 0.0972	4	− 0.0221	38
.08	0.0877	20	0.0968	6	0.0259	37
.09	0.0857	19	0.0962	6	0.0296	35
2.10	− 0.0838	19	+ 0.0956	7	− 0.0331	33
.11	− 0.0819	19	+ 0.0949	8	− 0.0364	32
.12	0.0800	19	0.0941	8	0.0396	28
.13	0.0781	18	0.0933	9	0.0424	28
.14	− 0.0763	18	+ 0.0924	9	− 0.0452	26
.15	0.0745	19	0.0915	10	0.0478	24
.16	0.0726	18	0.0905	10	0.0502	23
.17	− 0.0708	17	+ 0.0895	11	− 0.0525	21
.18	0.0691	18	0.0884	11	0.0546	20
.19	0.0673	17	0.0873	12	0.0566	18
2.20	− 0.0656	18	+ 0.0861	11	− 0.0584	17
.21	− 0.0638	16	+ 0.0850	13	− 0.0601	15
.22	0.0622	17	0.0837	12	0.0616	14
.23	0.0605	16	0.0825	13	0.0630	12
.24	− 0.0589	16	+ 0.0812	13	− 0.0642	11
.25	0.0573	16	0.0799	13	0.0653	10
.26	0.0557	16	0.0786	13	0.0663	9
.27	− 0.0541	15	+ 0.0773	14	− 0.0672	8
.28	0.0526	15	0.0759	13	0.0680	6
.29	0.0511	15	0.0746	14	0.0686	5
2.30	− 0.0496	15	+ 0.0732	14	− 0.0691	5
.31	− 0.0481	14	+ 0.0718	14	− 0.0696	3
.32	0.0467	14	0.0704	14	0.0699	2
.33	0.0453	13	0.0690	14	0.0701	2
.34	− 0.0440	14	+ 0.0676	14	− 0.0703	0
.35	0.0426	13	0.0662	14	0.0703	0
.36	0.0413	13	0.0648	14	0.0703	1
.37	− 0.0400	12	+ 0.0634	14	− 0.0702	2
.38	0.0388	12	0.0620	14	0.0700	3
.39	0.0376	12	0.0606	14	0.0697	3
2.40	− 0.0364	12	+ 0.0592	14	− 0.0694	4
.41	− 0.0352	12	+ 0.0578	14	− 0.0690	5
.42	0.0340	11	0.0564	13	0.0685	5
.43	0.0329	11	0.0551	14	0.0680	6
.44	− 0.0318	10	+ 0.0537	13	− 0.0674	6
.45	0.0308	11	0.0524	14	0.0668	7
.46	0.0297	10	0.0510	13	0.0661	7
.47	− 0.0287	9	+ 0.0497	13	− 0.0654	8
.48	0.0278	10	0.0484	13	0.0646	8
.49	0.0268	9	0.0471	12	0.0638	8
2.50	− 0.0259		+ 0.0459		− 0.0630	

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$		$\Phi(x)_5:16$		$\Phi(x)_6:32$	
2.50	+ 0.0022	— 0.0054	+ 0.0125	— 0.0259	9	+ 0.0459	13	— 0.0630	9
.51	+ 0.0021	— 0.0052	+ 0.0120	— 0.0250	9	+ 0.0446	12	— 0.0621	9
.52	0.0020	0.0050	0.0115	0.0241	9	0.0434	12	0.0612	9
.53	0.0019	0.0047	0.0111	0.0232	8	0.0422	12	0.0603	10
.54	+ 0.0018	— 0.0045	+ 0.0106	— 0.0224	8	+ 0.0410	12	— 0.0593	10
.55	0.0017	0.0043	0.0102	0.0216	8	0.0398	11	0.0583	10
.56	0.0016	0.0041	0.0097	0.0208	8	0.0387	12	0.0573	10
.57	+ 0.0015	— 0.0039	+ 0.0093	— 0.0200	7	+ 0.0375	11	— 0.0563	10
.58	0.0015	0.0037	0.0089	0.0193	7	0.0364	11	0.0553	10
.59	0.0014	0.0036	0.0086	0.0186	7	0.0353	11	0.0543	11
2.60	+ 0.0013	— 0.0034	+ 0.0082	— 0.0179	7	+ 0.0342	10	— 0.0532	10
.61	+ 0.0012	— 0.0032	+ 0.0078	— 0.0172	6	+ 0.0332	11	— 0.0522	11
.62	0.0012	0.0031	0.0075	0.0166	7	0.0321	10	0.0511	11
.63	0.0011	0.0029	0.0072	0.0159	6	0.0311	10	0.0500	11
.64	+ 0.0011	— 0.0028	+ 0.0069	— 0.0153	6	+ 0.0301	9	— 0.0489	10
.65	0.0010	0.0027	0.0066	0.0147	6	0.0292	10	0.0479	11
.66	0.0010	0.0025	0.0063	0.0141	5	0.0282	9	0.0468	11
.67	+ 0.0009	— 0.0024	+ 0.0060	— 0.0136	5	+ 0.0273	9	— 0.0457	11
.68	0.0009	0.0023	0.0057	0.0131	6	0.0264	9	0.0446	10
.69	0.0008	0.0022	0.0055	0.0125	5	0.0255	8	0.0436	11
2.70	+ 0.0008	— 0.0021	+ 0.0052	— 0.0120	4	+ 0.0247	9	— 0.0425	11
.71	+ 0.0007	— 0.0020	+ 0.0050	— 0.0116	5	+ 0.0238	8	— 0.0414	10
.72	0.0007	0.0019	0.0048	0.0111	5	0.0230	8	0.0404	11
.73	0.0007	0.0018	0.0045	0.0106	4	0.0222	8	0.0393	10
.74	+ 0.0006	— 0.0017	+ 0.0043	— 0.0102	4	+ 0.0214	7	— 0.0383	10
.75	0.0006	0.0016	0.0041	0.0098	4	0.0207	8	0.0373	10
.76	0.0006	0.0015	0.0039	0.0094	4	0.0199	7	0.0363	10
.77	+ 0.0005	— 0.0015	+ 0.0038	— 0.0090	4	+ 0.0192	7	— 0.0353	10
.78	0.0005	0.0014	0.0036	0.0086	4	0.0185	7	0.0343	10
.79	0.0005	0.0013	0.0034	0.0082	3	0.0178	6	0.0333	9
2.80	+ 0.0004	— 0.0012	+ 0.0033	— 0.0079	4	+ 0.0172	6	— 0.0324	10
.81	+ 0.0004	— 0.0012	+ 0.0031	— 0.0075	3	+ 0.0166	7	— 0.0314	9
.82	0.0004	0.0011	0.0030	0.0072	3	0.0159	6	0.0305	9
.83	0.0004	0.0011	0.0028	0.0069	3	0.0153	6	0.0296	9
.84	+ 0.0004	— 0.0010	+ 0.0027	— 0.0066	3	+ 0.0147	5	— 0.0287	9
.85	0.0003	0.0010	0.0026	0.0063	3	0.0142	6	0.0278	9
.86	0.0003	0.0009	0.0024	0.0060	2	0.0136	5	0.0269	8
.87	+ 0.0003	— 0.0009	+ 0.0023	— 0.0058	3	+ 0.0131	5	— 0.0261	9
.88	0.0003	0.0008	0.0022	0.0055	2	0.0126	5	0.0252	8
.89	0.0003	0.0008	0.0021	0.0053	3	0.0121	5	0.0244	8
2.90	+ 0.0003	— 0.0007	+ 0.0020	— 0.0050	2	+ 0.0116	4	— 0.0236	8
.91	+ 0.0002	— 0.0007	+ 0.0019	— 0.0048	2	+ 0.0112	5	— 0.0228	7
.92	0.0002	0.0007	0.0018	0.0046	2	0.0107	4	0.0221	8
.93	0.0002	0.0006	0.0017	0.0044	2	0.0103	4	0.0213	7
.94	+ 0.0002	— 0.0006	+ 0.0016	— 0.0042	2	+ 0.0099	5	— 0.0206	7
.95	0.0002	0.0006	0.0015	0.0040	2	0.0094	3	0.0199	7
.96	0.0002	0.0005	0.0015	0.0038	2	0.0091	4	0.0192	7
.97	+ 0.0002	— 0.0005	+ 0.0014	— 0.0036	1	+ 0.0087	4	— 0.0185	6
.98	0.0002	0.0005	0.0013	0.0035	2	0.0083	3	0.0179	7
.99	0.0001	0.0004	0.0012	0.0033	2	0.0080	4	0.0172	6
3.00	+ 0.0001	— 0.0004	+ 0.0012	— 0.0031		+ 0.0076		— 0.0166	

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$	$\Phi(x)_5:16$	$\Phi(x)_6:32$
3.00	+ 0.0001	— 0.0004	+ 0.0012	— 0.0031	+ 0.0076	— 0.0166
.01	+ 0.0001	— 0.0004	+ 0.0011	— 0.0030	+ 0.0073	— 0.0160
.02	0.0001	0.0004	0.0011	0.0028	0.0070	0.0154
.03	0.0001	0.0004	0.0010	0.0027	0.0067	0.0148
.04	+ 0.0001	— 0.0003	+ 0.0010	— 0.0026	+ 0.0064	— 0.0143
.05	0.0001	0.0003	0.0009	0.0024	0.0061	0.0137
.06	0.0001	0.0003	0.0009	0.0023	0.0058	0.0132
.07	+ 0.0001	— 0.0003	+ 0.0008	— 0.0022	+ 0.0056	— 0.0127
.08	0.0001	0.0003	0.0008	0.0021	0.0053	0.0122
.09	0.0001	0.0002	0.0007	0.0020	0.0051	0.0117
3.10	+ 0.0001	— 0.0002	+ 0.0007	— 0.0019	+ 0.0049	— 0.0113
.11	+ 0.0001	— 0.0002	+ 0.0007	— 0.0018	+ 0.0046	— 0.0108
.12	0.0001	0.0002	0.0006	0.0017	0.0044	0.0104
.13	0.0001	0.0002	0.0006	0.0016	0.0042	0.0100
.14	+ 0.0001	— 0.0002	+ 0.0006	— 0.0015	+ 0.0040	— 0.0096
.15	0.0001	0.0002	0.0005	0.0015	0.0038	0.0092
.16	0.0001	0.0002	0.0005	0.0014	0.0037	0.0088
.17	+ 0.0000	— 0.0002	+ 0.0005	— 0.0013	+ 0.0035	— 0.0084
.18		0.0001	0.0004	0.0013	0.0033	0.0081
.19		0.0001	0.0004	0.0012	0.0032	0.0077
3.20		— 0.0001	+ 0.0004	— 0.0011	+ 0.0030	— 0.0074
.21		— 0.0001	+ 0.0004	— 0.0011	+ 0.0029	— 0.0071
.22		0.0001	0.0003	0.0010	0.0027	0.0068
.23		0.0001	0.0003	0.0010	0.0026	0.0065
.24		— 0.0001	+ 0.0003	— 0.0009	+ 0.0025	— 0.0062
.25		0.0001	0.0003	0.0009	0.0024	0.0059
.26		0.0001	0.0003	0.0008	0.0022	0.0057
.27		— 0.0001	+ 0.0003	— 0.0008	+ 0.0021	— 0.0054
.28		0.0001	0.0002	0.0007	0.0020	0.0052
.29		0.0001	0.0002	0.0007	0.0019	0.0049
3.30		— 0.0001	+ 0.0002	— 0.0007	+ 0.0018	— 0.0047
.31		— 0.0001	+ 0.0002	— 0.0006	+ 0.0017	— 0.0045
.32		0.0001	0.0002	0.0006	0.0016	0.0043
.33		0.0001	0.0002	0.0006	0.0016	0.0041
.34		— 0.0001	+ 0.0002	— 0.0005	+ 0.0015	— 0.0039
.35		0.0001	0.0002	0.0005	0.0014	0.0037
.36		0.0000	0.0002	0.0005	0.0013	0.0035
.37			+ 0.0001	— 0.0004	+ 0.0013	— 0.0034
.38			0.0001	0.0004	0.0012	0.0032
.39			0.0001	0.0004	0.0011	0.0031
3.40			+ 0.0001	— 0.0004	+ 0.0011	— 0.0029
.41			+ 0.0001	— 0.0003	+ 0.0010	— 0.0028
.42			0.0001	0.0003	0.0010	0.0026
.43			0.0001	0.0003	0.0009	0.0025
.44			+ 0.0001	— 0.0003	+ 0.0009	— 0.0024
.45			0.0001	0.0003	0.0008	0.0023
.46			0.0001	0.0003	0.0008	0.0022
.47			+ 0.0001	— 0.0002	+ 0.0007	— 0.0020
.48			0.0001	0.0002	0.0007	0.0019
.49			0.0001	0.0002	0.0007	0.0018
3.50			+ 0.0001	— 0.0002	+ 0.0006	— 0.0017

Tafel für die Ableitungen von $\Phi(x)$.

x	$\Phi(x)_1$	$\Phi(x)_2:2$	$\Phi(x)_3:4$	$\Phi(x)_4:8$	$\Phi(x)_5:16$	$\Phi(x)_6:32$
3.50			+ 0.0001	− 0.0002	+ 0.0006	− 0.0017
.51			+ 0.0001	− 0.0002	+ 0.0006	− 0.0017
.52			0.0001	0.0002	0.0005	0.0016
.53			0.0001	0.0002	0.0005	0.0015
.54			0.0000	− 0.0002	+ 0.0005	− 0.0014
.55				0.0001	0.0005	0.0013
.56				0.0001	0.0004	0.0013
.57				− 0.0001	+ 0.0004	− 0.0012
.58				0.0001	0.0004	0.0011
.59				0.0001	0.0004	0.0011
3.60				− 0.0001	+ 0.0003	− 0.0010
.61				− 0.0001	+ 0.0003	− 0.0010
.62				0.0001	0.0003	0.0009
.63				0.0001	0.0003	0.0009
.64				− 0.0001	+ 0.0003	− 0.0008
.65				0.0001	0.0003	0.0008
.66				0.0001	0.0002	0.0007
.67				− 0.0001	+ 0.0002	− 0.0007
.68				0.0001	0.0002	0.0007
.69				0.0001	0.0002	0.0006
3.70				− 0.0001	+ 0.0002	− 0.0006
.71				− 0.0001	+ 0.0002	− 0.0005
.72				0.0001	0.0002	0.0005
.73				0.0000	0.0002	0.0005
.74					+ 0.0001	− 0.0005
.75					0.0001	0.0004
.76					0.0001	0.0004
.77					+ 0.0001	− 0.0004
.78					0.0001	0.0004
.79					0.0001	0.0003
3.80					+ 0.0001	− 0.0003
.81					+ 0.0001	− 0.0003
.82					0.0001	0.0003
.83					0.0001	0.0003
.84					+ 0.0001	− 0.0003
.85					0.0001	0.0002
.86					0.0001	0.0002
.87					+ 0.0001	− 0.0002
.88					0.0001	0.0002
.89					0.0001	0.0002
3.90					+ 0.0001	− 0.0002
.91					0.0000	− 0.0002
.92						0.0002
.93						0.0001
.94						− 0.0001
.95						0.0001
.96						0.0001
.97						− 0.0001
.98						0.0001
.99						0.0001
4.00						− 0.0001

H Bruns Wahrscheinlichkeits-
rechnung und Kollektivmasslehre. L. u. B. 1906.

At give en Lærebog om et mathematisk Emne den historiske Form kan være heldigt. Men Betingelsen er at Emnets Historie er løbet ud, saa at Forfatteren overlegent kan føre Læserne udenom alle Historiens Omveje og blinde Gyder. Staar Forf. selv midt i Bevægelsen, især hvor denne nærmer sig en Krise, kan det være ham umuligt at opnaa systematisk Klarhed, det vil i alt Fald være ham lettest at referere Forgængernes Arbejder, men det er da ikke nok at han kritiserer en Del af deres Vildfarelser, der bliver let saa meget tilbage at hans Bog bliver uanvendelig for Begyndere eller dog lader dem blive staaende længere tilbage i Bevægelsen end Forfatteren selv var naaet.

Dette gælder i høj Grad om Iagttagelseslæren, som fra Sandsynlighedsregningens Spilleproblemer gennem Omveje er naaet ~~[illegible]~~ til den Krise som fremkaldtes ved den Efter-Darwinske Biologi og Fechners Kollektivmasslehre m. m. Prof. Bruns er med sin ovennævnte Bog naaet helt hen foran Gennembruddet, men er standset foran Springet, og som Lærebog er hans Værk derfor forfejlet. For danske Læsere,

som med Tilknytning til det nye kan
ønske at sætte sig ind i de gamle Synsmaader
vil Værket ~~[illegible]~~ dog have megen Interesse,
om det end heller ikke for dem vil være let at læse.

Forf. er naaet saa vidt at han ser
Forskellen mellem faktiske og præsumptive
Fordelingslove (eller Fejllove). Men han gennem-
fører den ikke. Han erkender Kollektivbegrebets
~~faktiske~~ oprindelig faktiske Karakter men
ved at forudsætte Gentagelsernes Antal stort
nok, gaar han udenom de faktiske Fordelingers
Afhængighed af Antallet. Han taler med
Forkærlighed om relative Hyppigheder (han skriver
r H K) ved Siden af Sandsynligheder, men for
ham er dette endnu et og samme Begreb.

For Forf. ~~[illegible]~~er Fordelingernes mathematiske
Udtryk bestandig Hyppighedsfunktionen og
noget andet. For fra den tabellariske Oversigt
over Gentagelserne at naa til Hyppigheds-
funktionen tvinges han vel til at anvende
baade Potenssummerne og mere hensigtsmæssig
symmetriske Funktioner, men han overser at
Kollektivgenstandens naturligste mathematiske
Definition er at slaa sammen som Rødder i
en Ligning af Antallets Grad, og at derfor deres
symmetriske Funktioner er særlig berettigede

som Udtryk for Fordelingen, og at de er overlegne over Hyppighedsfunktionerne, fordi de er uafhængige af Hypothese om dennes Form.

Ogsaa i en anden Henseende viser B sig Konservativ. Overfor K.G søger han at finde deres alm. analytiske Udtryk, og at aflede dettes Form ved mathem. Spekulation. Ogsaa her bringes han dog til Indrømmelser. Den enkelte K.G. finder sit Udtryk ikke i den anal. Funktions Form men i Konstanterne i dennes Led, og disse Konstanter maa fremstilles som Funktioner af de enkelte givne Iagttagelser. Han undgaar at fremstille dem som symmetriske Funktioner eller som Koeff. i Ligningen af mte Grad. Men maa benytte Potenssummerne af alle de lavere Ordener til Beregning af Elementerne i den af ham foretrukne Normalrække. Thi vel kan han komme til disse ved Grams Løsning (Analogi med Fouriers Rækken), men til virkelig Regning maa Potenssummerne i alt Fald som Gjennemgangsled. Han søger yderligere at lette Regningen ved Operationer der er identiske med Oppermanns Transformation fra 1871. Men disse Systemer af symm. Funktioner er for B kun Mellemled; som selvstændigt Udtryk vil han kun anvende sin Normalform.

Er B's Normalform da et Universalmiddel
B søger at godtgøre dette ved pinlige Konvergens
undersøgelser som kun delvis gennemføres. Men
her spørges der ikke blot om Konvergens; fuldt
saa vigtigt er det at en god analytisk Frem-
stilling af Hyppighederne ikke bør blive negativ.
Men denne Betingelse ophæver Uafhængigheden
mellem de enkelte Elementer i B's Normalformen
og stiller den i Klasse med Halv-invarianterne
og andre symm. F. Dette overser Bs. ganske.

Tanken om et Universaludtryk for R.G.
maa for Tiden helt opgives. Der maa gives
Valg imellem talrige Udtryksformer, og
aabnes som fra den ene kan føre til den anden
eller bedst fra en Centralstation til alle de
andre. Som saadan er det jeg anbefaler μ_n

Ved heller at gaa udenom (?) skader B. sin Fremstilling paa flere Punkter (?) Ers. de kunne Funktionærs Fordelingsloven. Og dette er maaske Grunden til at han erklærer Udbetaling for K.M.L. uvedkommende.

Sandsynlighedsregningen gennemgaar i vor Dage en Krisis.

Laplace Gauss
Bertrand?

Bessel – Darwin, Statistiken, Fechner

Statistikernes og Biologernes Krav og deres halvstuderede Forsøg og vilde [illegible] gør Krisen akut, og kræver at mulig Anstrengelse ikke blot ~~for~~ rettet paa at naa det korrekte Maal, men ogsaa ~~at~~ et hensynsfuldt Valg af Midler.

Anmeldtens Ret er at han har forsøgt en radikal, systematisk Løsning af Opgaven.

I den anm. Bog er B. ingenlunde radikal, men konservativ. Under sine Bestræbelser for at bevare saa meget som muligt, og at gjøre grundige Behandlinger af Spørgsmaalene, som Fechner [illegible] [illegible] bringes den højt ansete Prof. dog ~~et~~ Skridt for Skridt til at

gaa med til Grundbrudet, men han bliver staaende der, og refererer sine mangeaarige Forelæsninger uden at bearbejde dem ud fra det i Grundbrudet ændrede Stade. Hans Bog er forholdsvis let for dem der er oplært med de gamle Lærdomme. Men den vil ikke blot være vanskelig at bruge men ogsaa synes usammenhængende for dem der vil blive oplært efter at Krisen er overstaaet.

Efter i de første Forelæsninger at have behandlet S.-R. saa temmelig i gammel Stil, og dertil knyttet nogle Afsnit af Integralregningen, behandler Prof. i 10nde Forel. Fechners Begreb K.G. med Veneration men dog med Kritik, som ikke altid er heldig. Fechner viser mere Respekt for Baand. Paa andre Punkter er det bedre, og B. definerer der K.G. som „eine Vielheit von gleichartigen Dingen, die nach einem veränderlichen (oder fleren) Merkmal zahlenmässig geordnet werden können."

Bruns begynder „Vielheden" til stort nu. for at kunne slutte om Regelmæssighed eller U.R. og overhovedet naa et „Ergebniss."

Hvad „Gleichartigkeit" angaar er B til Begyndelse mere liberal end jeg, men efterhaanden trækker han ogsaa her af Hensyn til Ergebniss saa meget fra; at han maa kunne slutte sig til mig naar jeg kalder K.G. Gentagelserne af samme Iagttagelse, naar ~~dette~~ Iagtt. tages i vedst. Forstand som jeg gør, ligger heri kun den Antagelse om god Mening og Kausalitet som vilde lade os vente et vist konstant Udfald, hvor de væsentlige Omstændigheder (Merkmale) er holdt konstante, medens vi faar flertydigt Svar som Følge af de som uvæsentlig ansete Omstændigheder (Tilfældigheder). Havde B. stillet sig paa dette Standpunkt vilde ogsaa han have lagt Vægt paa K.G. som Rødderne i en Lign af n^te Grad

I 15. Forelæsning foregaar Gennembrudet. Det er jo ikke nok at Antallene er store, til at sikre „Ergebniss" maa ogsaa vides hvor stor Usikkerhed det er. Her spørges om Fordelingslovenes Fordelingslove og Afhængighed af Grundlagenes Antal. Overfor denne virkelig betydelige Vanskelighed tvinges B til at opstille Forskel mellem „Sth" og „Sd" overfor K-G. Men dette er netop Forskellen mellem faktiske og præsumptive Fordelingslove. Dermed er B. traadt ud af det gamle og ind i det nye. Men dette burde have præget Bogen fra først af. jvf. pag 4. Forskel mellem rth. og W, og Bortfalden i Stedet for Kritik af Bayes' Princip med Opfattelse af W(F) som virkelige Sandsynligheder. Af historiske Grunde at fastholde en Fejl er altfor konservativt. I 10. Forelæsning osv. vides da K-G. have faaet den Begrænsning det trænger til.

Pg 2 Intet væsentligt nyt i W. R. siden Laplace, kun mere omfattende Opgaver — Fechner

Pg 11 v. K = W. (NB ikke gunstig).

Pg 14 Nödvendigheden af en empirisk Bekræftelse af Loven om de store Tal indses, men dennegang vil B. ikke lægge det faktiske til Basis

Pg 23. Brüns savner et Ord for „væsentlige Omstændighed.

Pg 24. Viser det sig hvorledes man föres ind paa det „Bayesske Princip" naar man ikke vil lægge Vægt i de faktiske Fordelingslove. W(H), Sandsynligheden for Hypothesens posteriori ensartet med det oprindelige Sandsynlighedsbegreb.

Pg 85 B. vil have „Fehlertheorie" Forsikrings- og Statistik gjort selvstændige og adskilt fra W. R.

Pg 86 I § 63 omtaler B. Juryproblemet paa en saadan Maade at det kunde fört ham ikke blot til et kört Regneexempel men henimod den faktiske Fejllov. Han ser nemlig med Rette at objektivt er Problemet ... ved Raad-Snakkemethoden — og drejet af.

Pg -90- § 63-67 „Lexis" B gaar med mange interessante Ord omkring det faktiske som Katten om den varme Gröd.

Pg 93 Fechner, § 68. „genügend genau"

Pg 95 i § 68 har Fechner fanet Ordet for faktiske Lov i § 69 omtaler B. Fechners mislykkede Technik som om ikke var hans væsentligste Fortjeneste

Pg 95 For Richard staar KML selvstændig ved Siden af W.R. Brauns siger at i den maade er in det KML ledigligh med det følgerechtlen Ausbau der Methoden des W.R zu tun hat.; altsaa W.R. som det primære; han burde have vendt det om og gjort den faktiske KML til det primære, og tværs imod hans Ord har vi set det Følgende ndk netop dette. Jvnf. for hans Ord om Bernoulli og Poisson's Sætninger som Følger af „theoretisk konstruerede Modi“ de a præsumptiv Regllove.

Pg 96 Om R.G fremhæver at de skal adskille sig ved et variabelt Maal, men der siges ikke at dette skal være det der interesserer os som formelt Virkning af de øvrige Omstændigheder, og skulde være konstant med dem, hos de udløsende Aarsagers Rækker.

Pg 96. Flerdimensionale R.G anerkendes, uagtet det formelt sprænger B's Definition, og det siges øvrigt om dem at deres Behandling kan føres tilbage til de endimensionale.

97 gross genug

Pg 172 B. anser principiell Udjævning for at ligge udenfor KML's Grænser

172. At paavise Baand i Indtegnelsesrækken lader sig ikke gøre ved Tæl lovs Regllovs Grænseværdier for reelle Sandsynligheder og positive Hyppighedér.

Pg 123 Bruns vil have den sandsynlige Fejls henvist til Pulterkammeret for historiske Alterlinnet, ja! + et og andet mere.

123 B. indfører Navnet numeriske Elementer for de symmetriske Funktioner han [illegible], men taler naturligvis ikke om at de er symmetriske Funktioner.

Pg 124 omtaler B. at Rækken $\Sigma\phi$... bliver endelig i en faktisk Fejllov, og beviser dette ved at aflede den af Potenssummerne, men at derved bestemmes af Ligningskoeff. som sym. Funktioner undgaar han omhyggeligt at nævne.

Pg 125 Kritiserer B Potenssummerne som Udtryk for Fordelingsloven og bemærker rigtigt, at de bør erstattes af andre (symm.) Funktioner af dem men han fejler ved at gøre dette gældende overfor Kubiksummerne, og hvor man venter en Forklaring angaaende Betydningen af D_3, D_4 bryder han pludselig af og indskyder en § 94. som angiver Indholdet af de næste 5 Kapitler

Pg 136 Ved F.L. for Funktioner af flere Variable kan B kun paa Lovens form for D_3, D_4, D_5 som udløb af Halvinvarianter, men gjort Afkald paa de højere, uden at anerkende deres Underlegenhed under H.S.

Pg 161 Sollværdi = presumptiv Værdi i Modsætning til Tal = = faktisk.

Pg 4 Tilløb til Sondring mellem fuldkt. og præs. F.

Pg 13 Stærk Tilnærmelse til Loven om de store Tal i den Form, jeg anvender.

Pg 200 Fejlloven i Sandsynlighed afhænger af det ene Tal p. – B. erkender Uligheden heraf.